INVESTIGATION OF NEW PROPERTIES AND APPLICATIONS OF QUADRUPLEX DNA AND DEVELOPMENT OF NOVEL OLIGONUCLEOTIDE-BASED TOPOISOMERASE I INHIBITORS

WANG YIFAN

NATIONAL UNIVERSITY OF SINGAPORE

2008

INVESTIGATION OF NEW PROPERTIES AND APPLICATIONS OF QUADRUPLEX DNA AND DEVELOPMENT OF NOVEL OLIGONUCLEOTIDE-BASED TOPOISOMERASE I INHIBITORS

WANG YIFAN

(B.Sc., Soochow University, China)

A THESIS SUBMITTED
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF CHEMISTRY

NATIONAL UNIVERSITY OF SINGAPORE

2008

Acknowledgements

I would like to express my wholehearted gratitude to my supervisor, Associate Professor Li Tianhu for his profound knowledge, invaluable guidance, constant support, inspiration and encouragement throughout my graduate studies. He is not only an extraordinary supervisor, a complete mentor, but a truly friend. The knowledge, both scientific and otherwise, that I accumulated under his supervision, will aid me greatly throughout my life.

I also give my sincere thanks to all the members of the Li group: Li Xinming, Li Ming, Liu Xiaoqian, Xu Wei, Magdeline Tao Tao Ng and Chua Sock Teng, for their cordiality and friendship. We had a great time working together.

I wish to express my deepest appreciation to my family and my boyfriend for their love and support. Without their help, I can not complete this work.

Last but not least, my acknowledgement goes to National University of Singapore for awarding me the research scholarship and for providing financial support to carry out the research work reported herein.

Table of Contents

Summary

Some sequences of DNA that possess certain guanine or cytosine-riched stretches are capable of associating into two types of four-stranded DNA structures, namely G-quadruplex and i-motif respectively. It has been suggested in the past that some of these quadruplex structures could exist in some biologically important regions of DNA such as at the end of chromosomes and in the regulatory regions of oncogenes. In addition, due to their distinctive structural characteristics, quadruplex structures of DNA have been widely used as building blocks in various nanotechnological applications, such as G-quadruplex nanodevices and i-motif nanoswitches. With the aim of exploring new properties and applications of quadruplex DNA during my graduate studies, we have (1) constructed i-motif DNA-based molecular devices that are operable through variations of their surrounding pH values; (2) developed certain fluorescence-tagged circular G-quadruplexes to be used as molecular probes; and (3) investigated the factors that affect the G-quadruplex that could undergo self-cleavage reactions. Finally, we have designed and synthesized certain dumbbell-shaped oligonucleotides and further examined their inhibitory effects on the activities of human topoisomerase I.

In Chapter 2, design and synthesis of a novel quadruplex DNA machine is presented that was capable of converting chemical energy into elastic potential energy. As a consequence of this energy converting process, Watson-Crick hydrogen bonding interaction between two complementary 11-mer oligonucleotides was forced to break down, leading to a free energy change of 12.46 kcal mol^{-1}.

In Chapter 3, self-cleavage reaction of a guanine-riched oligonucleotide was thoroughly studied during our investigation. Subsequent examinations on certain factors that affect self-cleavage reactions of G-quadruplexes are described, such as

variation of metal ions, pH values and concentration of DNA. In addition, kinetic analysis of self-cleavage of G-quadruplex was also carried out. It is our hope that the results reported in this chapter could be helpful for searching for new G-quadruplex structures that could perform self-cleavage reactions.

In Chapter 4, our studies of synthesis and characterization of unimolecularly circular G-quadruplex on the template basis of G-quadruplex through chemical ligations of guanine-riched oligonucleotides are described. Loop-size effect of ligation reaction, conformation dependence of circularization course, effects of alkali ions and pH values as well as concentration of potassium ions on the circularization reactions were investigated during our studies. The potential application of the obtained unimolecularly circular G-quadruplex in certain biological processes is also presented in this chapter.

In Chapter 5, design and synthesis of a series of dumbbell-shaped circular oligonucleotides containing internal C3-spacers are presented. Our studies demonstrated that this C3-spacer-containing oligonucleotide displays an IC_{50} value of 33 nM in its inhibition on the activity of human topoisomerase I, which is much efficient than those of camptothecins (anticancer drugs currently in clinical use).

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Basic Information about DNA

Deoxyribonucleic acid (DNA) is a type of biomacromolecule that contains genetic information used for the functioning of living organisms.[1] The major role of DNA *in vivo* is its long-term storage of genetic information. From the perspective of chemistry, DNA is a long polymer built up on simple units called nucleotides, linked together through a backbone made of sugars and phosphate groups.[1, 2] A single strand form of DNA is a long chain composed of different nucleotides. Each nucleotide consists of a sugar, a phosphate and a nitrogenous base. There are four different types of bases in DNA (**Figure 1-1**), and each base is usually abbreviated by the first letter of its name: Adenine (A), Thymine (T), Guanine (G) and Cytosine(C). Two strands of nucleotides usually wrap around each other, which are twisted together into a long helix; like a ladder twisted about its long axis (**Figure 1-2**).[2] The backbone of sugar-phosphate linkages forms the uprights of the twisted ladder. The rungs of the ladder are made up of base pairs, which are almost always found connected to each other. Each twist of the ladder contains approximately 10 rungs, which is 0.34 nm apart. In a complete helix, A always lines up with T and G goes with C. In these combinations, the different bases fit together perfectly like a lock and key, which is termed with "Watson-Crick base pairing" (**Figure 1-2**).[2]

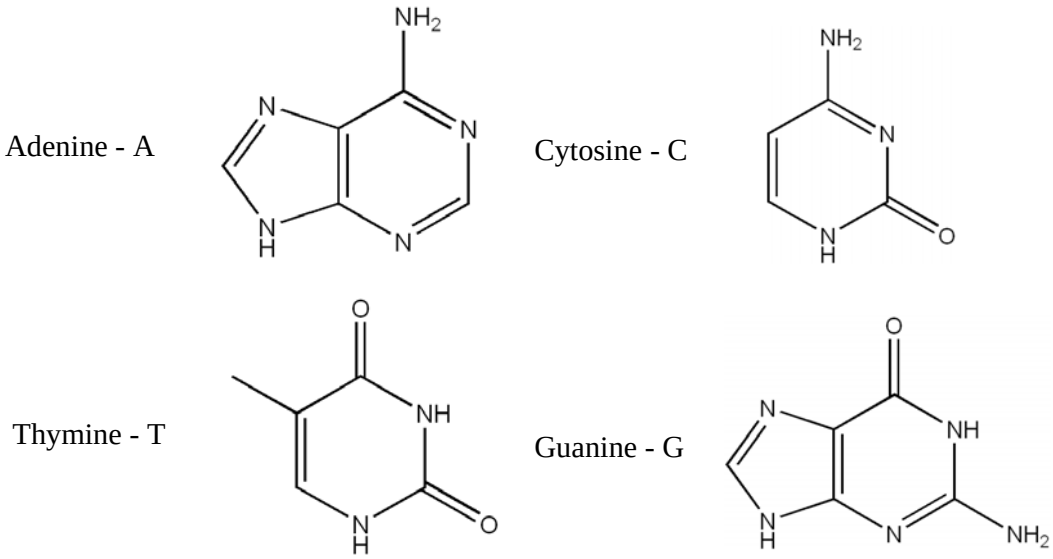


Figure 1-1. Structures of four types of nitrogenous bases

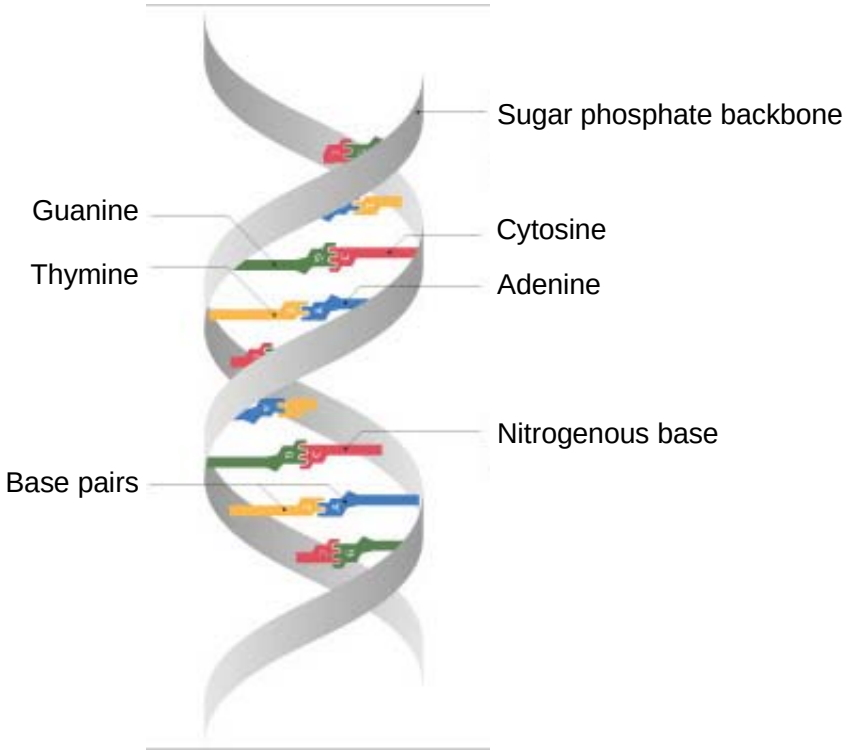


Figure 1-2. Base Pairing in DNA Double Helix

1.2 G-Quadruplex Form of DNA

1.2.1 Guanine Quartets

DNA commonly exists in the form of duplex structure in which two self-complementary strands are held together by Watson–Crick base pairs. Besides this form of duplex DNA, certain guanine-riched DNA sequences can form four-stranded structures, namely G-quadruplexes.[3-6] The basic building block of G-quadruplex is the guanine quartets (also known as guanine tetrads) composed of four guanine bases arrayed in a square planar configuration, in which each base is both the donor and acceptor of two hydrogen bonds with its neighbors (**Figure 1-3**). More precisely, the guanine quartet arises from the association of four guanines into a cyclic Hoogsteen hydrogen bonding arrangement that involves N1, N7, O6 and N2 of each guanine base. [7-10] Positively charged metal ions can be sandwiched between the quartets. Their presence in the central cavity of the quadruplex helps maintain the stability of the tetraplex structure.[3] In addition, the G-quartet could form a particularly effective stacking unit when placed next to each other, resulting in a strong attraction that contributes substantially to the stability of the overall structure.[11-19]

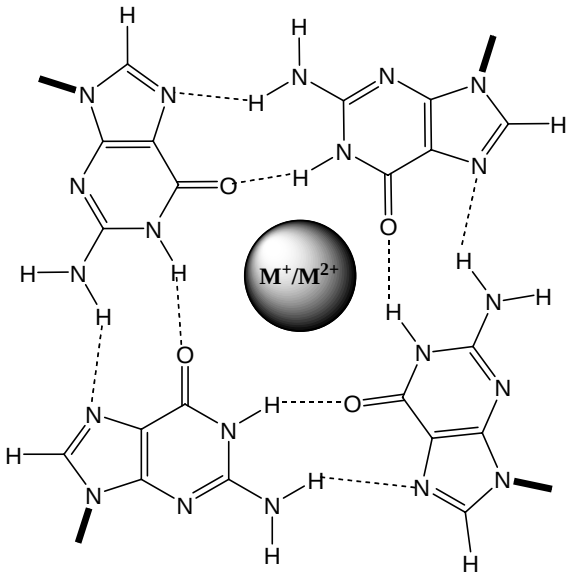


Figure 1-3. Structures of Guanine Quartets

1.2.2 G-quadruplexes

In certain guanine-riched strands, two or more G-quartets can stack upon each other to form four-stranded structures with a guanine tetrad core. These structures are known as **G-quadruplexes**.[3] The term G-quadruplex refers to any four-stranded DNA structure containing guanine quartets without reference to strand connectivity. [3-5]

G-quadruplexes exhibit an unusual dependence on specific metal ions, usually K^+ and occasionally Na^+,[20] which results in very tight metal binding via inner sphere coordination. The cavity between G-quartets is well suited to coordinating the right size of cations because the two planes of quartets are lined by eight carboxyl O6 atoms from guanine. It was reported that a wide variety of cations are capable of occupying the central cavity of quadruplex structures, including monovalent ions such as NH_4^+ and Tl^+ and divalent cations such as Sr^{2+}, Ba^{2+}, and Pb^{2+}.[21]

1.2.2.1 Discovery of G-quadruplex DNA

It was known since early 19th century that guanosine and its derivatives could form viscous gels in water.[22] Until 1962, David R. Davies *et. al.*[23] proposed on the basis of X-ray diffraction data that four guanine bases form a planar structure through Hoogsteen hydrogen bonding interaction.[22] Subsequent NMR studies of these gels further suggested that cations such as Na^+ and K^+ could coordinate to the O6 atoms of each guanine base and strongly influence the specific type of structure adopted by the gels.[24]

1.2.2.2 Structural Polymorphism of G-quadruplex Structures

One of the most intriguing aspects of G-quadruplex is their extensive polymorphism which arises from variation of strand stoichiometry, strand polarity and connecting loop.[11-15] Quadruplexes typically contain 1, 2, or 4 nucleic acid strands,

giving rise to unimolecular, bimolecular or four-stranded structures and display a wide variety of topologies (**Figure 1-4**). These tetraplex structures can exist in different isomeric forms caused by different strand polarities of adjacent backbones. Certain guanine-riched sequences can, for example, orient themselves in all parallel, three parallel and one anti-parallel, adjacent parallel or alternating anti-parallel.[10] Some of the polymorphisms are discussed in the following sections.

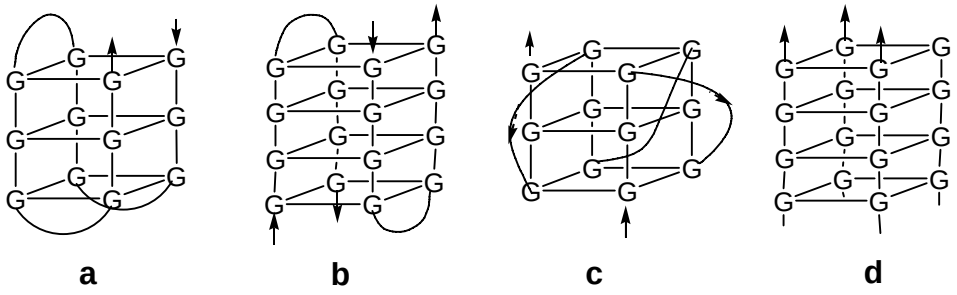


Figure 1-4. G-quadruplex structures formed from one, two or four strands

1.2.2.2.1 Strand Stoichiometry

G-quadruplexes could be formed by association of one (**Figure 1-5A**), two (**Figure 1-5B**), or four strands (**Figure 1-5C**) of oligonucleotides. The structural assemblies of unimolecular, biomolecular and tetramolecular G-quadruplexes could display different physical and chemical properties.

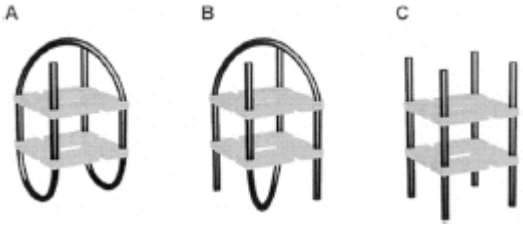


Figure 1-5. Stoichiometries of G-Quadruplex structures

1.2.2.2.2 Strand Polarity Polymorphism

The additional structural characteristic of G-quadruplex is the relative arrangement of adjacent backbones, which could have different polarities. The four

strands of oligonucleotides in a G-quadruplex can be all parallel (**Figure 1-6A**), three parallel and one anti-parallel (**Figure 1-6B**), adjacent parallel (**Figure 1-6C**), or alternating anti-parallel (**Figure 1-6D**). Many guanine-riched oligonucleotides have been determined either with NMR[26] or crystallography[27], which displayed different strand polarities as shown in **Figure 1-6**.

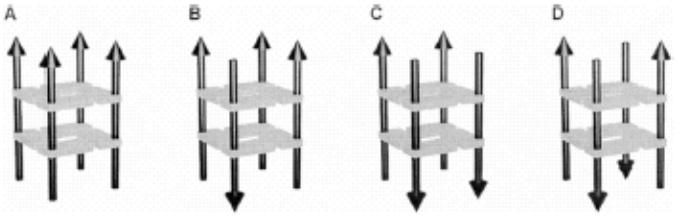


Figure 1-6. Different strand polarity arrangements of G-quadruplexes

1.2.2.2.3 Connecting Loops

The loops that connect guanine quartets participating in the formation of unimolecular or bimolecular G-quadruplexes can run in different ways. The two strands involved in bimolecular G-quadruplexes can have loops that connect guanine tracts either diagonally or edgewise. [25]

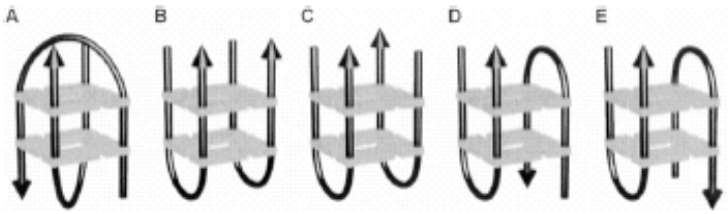


Figure 1-7. Strand connectivity alternatives for bimolecular guanine tetrad structures

Diagonal loops are expected to protrude on opposite ends of the guanine tetrad core (**Figure 1-7A**). When the two loops connect guanine tracts edgewise, they can be either on the same or on opposite sides of the tetrad core. Loops on the same side of the core can be either parallel (**Figure 1-7B**) or anti-parallel (**Figure 1-7C**). When the

two loops protrude on opposite sides of the core, they can run in two different directions (**Figure 1-7D** and **1-7E**).

For unimolecular G-quadruplexes, structural isomers of G-quadruplex caused by loop-connecting fashion are fewer. In order to avoid the clash of two diagonal loops on the same side, the three loops can join either in the order adjacent-adjacent-adjacent (**Figure 1-8A**) or adjacent- diagonal-adjacent (**Figure 1-8B**). On the other hand, there are some examples of parallel strands connecting via loops running on the outside of the guanine tetrad core (**Figure 1-8C**), which indicates that the spectra of unimolecular structures may be more complex than prospected here.[28]

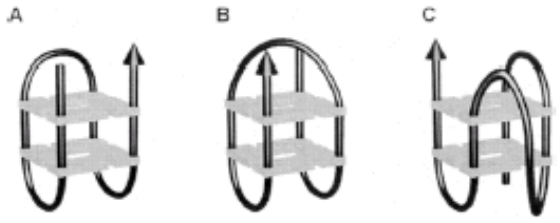


Figure 1-8. Strand connectivity alternatives for unimolecular guanine tetrad structures

1.2.2.3 Possible Roles of G-quadruplex *in vivo*

Little attention was paid to the phenomenon of guanine tetrads for more than 20 years since it was elucidated in 1962 by David R. Davies. Until 1980s, emerging interest in G-quadruplex structure was stimulated by several implications of its existence in various biologically important genomic regions such as telomeres.[29, 30] For example, these structures were suggested to participate in telomere regulations. In addition, it is believed that G-quadruplex is responsible for the switch recombination to bring different constant regions next to variable regions during the differentiation of B lymphocytes.[31]

In addition, telomeres are the specialized ends of linear chromosomes comprising tandemly-repeated short DNA sequences.[32] Various proteins are involved in regulating the structure and function of human telomeres, including telomerase and some telomere-interacting proteins such as Pot1, TRF1 and TRF2. It is well known that telomeres are essential for genome integrity and appear to play an important role in cellular aging and cancer. In almost all organisms, the telomeric DNA sequence has a G-rich 3' overhang, such as "TTAGGG" in vertebrates or "TTGGGG" in ciliate Tetrahymena. The length of the sequences can range from a dozens to thousands of such repeats. Generally, the last few hundred based of G-rich strand in telomeres is thought to be in single-stranded form.[32] Besides present at the ends of telomeres, guanine-rich sequences are found in a number of important DNA regions, such as in the immunoglobulin switch regions and gene promoter region of c-myc and other oncogenes.[36] Moreover, several G-quadruplex-binding proteins have been identified over the past 10 years.[32-35] It consequently becomes apparent that G-quadruplex could play certain significant roles in various types of biological processes.

1.2.2.3.1 G-quadruplex-Interactive Proteins

Many proteins, mostly from ciliates and yeast, have been found to bind to G-quadruplex structures.[32-40] Among these, yeast RAP1 protein[34] and beta-subunit of Oxytricha telomere binding protein[37] are the most interesting ones because they not only bind to G-quadruplex but also facilitate the formation of these structures. In addition, four helicases, the Simian Virus (SV) 40 large T-antigen,[41] Bloom's syndrome helicase (BLM) from yeast, and Werner syndrome helicase from humans[42] have been found to unwind G-quadruplex DNA. Another enzyme that could interact with quadruplex structures is human DNA topoisomerase I (**Topo I**).[43] Arimondo *et. al.* demonstrated that Topo I can bind to both linear, four-stranded quadruplexes and

unimolecular quadruplexes. Moreover, it was demonstrated that this enzyme can induce the formation of four-stranded G-quartet structures. [43]

1.2.2.3.2 Telomere Protection and Elongation

Telomeric proteins are known to bind both double-helical telomeric DNA as well as single-stranded, non-quadruplexed telomeric DNA. Zahler *et. al*[44] illustrated that the folded quadruplex form of the 3' telomere overhang is a poor substrate for telomerase, and accordingly proposed that quadruplex formation may play a role in the negative regulation of telomerase-based replication.[45-46]

Formation of G-quadruplex structure to afford 3' overhang protection has been proposed as the molecular mechanism for telomere protection.[47] It was suggested that the 3' overhang could fold over to form an intramolecular G-quadruplex. Such structures are most likely to form during replication when long single-stranded G-rich tails are expected to be transiently present[47] (**Figure 1-9A**). It was further demonstrated that certain single-stranded G-rich overhang might fold back to form a hairpin structure involving G-G base pairing (**Figure 1-9B**).[46] Two such hairpins from different chromosomes can then dimerize to form a G-quadruplex structure and help the alignment of sister chromatins.

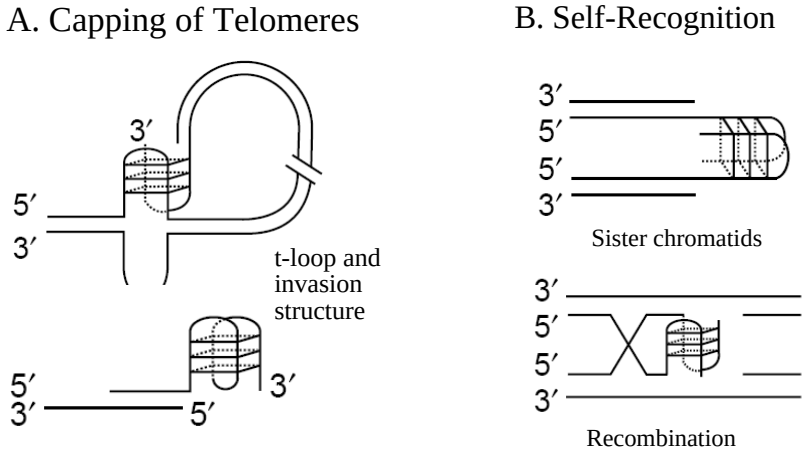


Figure 1-9. Possible biological roles of G-quadruplexes

1.2.2.3.3 Interaction of Small-Molecule with G-Quadruplex

Zahler *et. al.* demonstrated in 1991 that K^+-stabilized G-quadruplex structures were able to inhibit telomerase activity.[48] Since then, G-quadruplex DNA has become an attractive target for design of telomerase inhibitors.[49] Several groups subsequently used structure-based design approach to develop lead compounds that interact with G-quadruplexes in order to inhibit telomerase activity and disrupt the function of telomeres.[50, 51] After the original discovery of G-quadruplex interactive telomerase inhibitors (*e.g.* anthraquinones), a number of compounds such as fluorenones, bi-substituted acridines and cationic porphyrins have been identified, and their interactions with G-quadruplex have also been studied extensively.[52-55]

1.3 i-Motif Structure of DNA

Besides G-quadruplex form of DNA, certain Cytosine-rich oligonucleotides could form tetraplex assemblies at low pH,[56] namely ***i-motif***. The structural entity is composed of two parallel-stranded DNA duplexes zipped together in an antiparallel orientation and held together by hemiprotonated $C{\cdot}C^+$ base pairs.[56-60] NMR studies showed that the same C-rich strands of oligodeoxynucleotides can form intercalated structures of i-motif that differ essentially in their intercalation and loop topologies.[61]

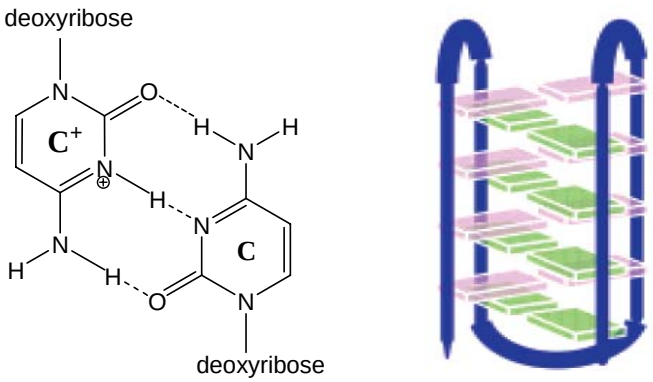


Figure 1-10. Illustration of $C{\cdot}C^+$ interaction in i-motif structure of DNA

1.3.1 Discovery of i-Motif Form of DNA

In 1993, Gehring *et. al* [56] demonstrated that certain oligomers containing tracts of cytidine could form hemiprotonated base pairs at acid pH. This research group solved the structure of d(TCCCCC) and found that it was a four-stranded complex in which two base-paired parallel-stranded duplexes were intimately associated and their base pairs were fully intercalated. Subsequent NMR analysis showed only six spin systems, indicating that the structure is highly symmetrical on the NMR timescale and the four strands are equivalent.[56, 61] The outcomes of these studies demonstrated that certain C-riched sequences could exist indeed in tetraplex forms of i-motif as illustrated in **Figure 1-10**.

1.3.2 Stoichiometries and Topologies of i-Motif DNA

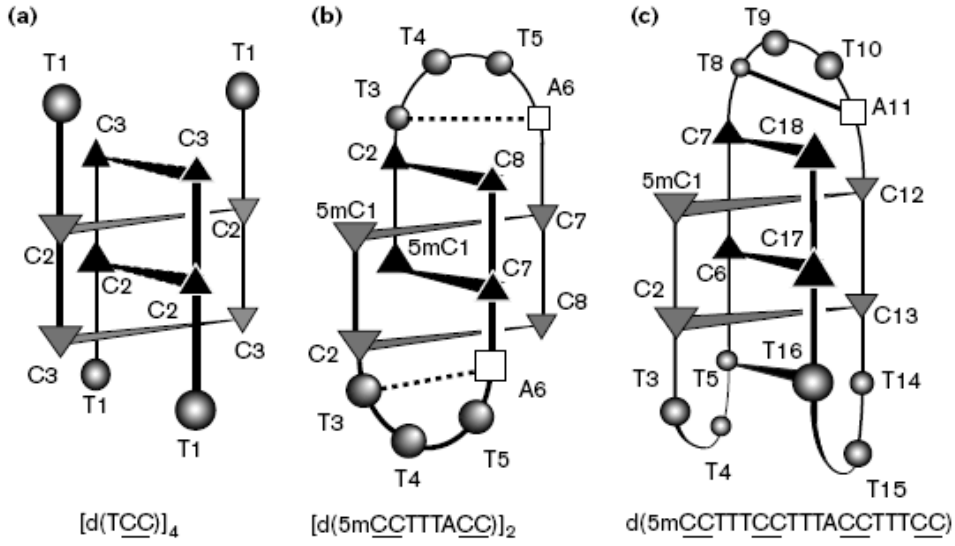


Figure 1-11. i-motif structures with (a) four, (b) two and (c) one strand(s)

i-Motif is the only known nucleic acid structure containing systematic base intercalation,[58] in which two cytidine stretches form a parallel-strand duplex via $C{\cdot}C^{+}$ base pairs and two such duplexes associate head-to-tail by base-pair intercalations into a quadruplex. (**Figure 1-11**) The intercalated structure is partially stabilized by

interactions of hydrogen-bonded protonated and neutral cytosines. Consequently, formation of i-motif structure generally require acidic environment in order to protonate one of the cytidines in $C{\cdot}C^+$ base pairs (pKa = 6.2).[57-60] It has been suggested that telomeric C-rich strands of tetrahymena and of vertebrates may fold into a monomeric i-motif that persisted at pH 7.0 despite the requirement for cytidine protonation. The structural feature of i-motif is accordingly expected to play certain biological roles *in vivo*.[62]

In addition, as shown in **Figure 1-11**, oligonucleotides such as d (Cn) can form two fully intercalated i-motif tetramers (**Figure 1-11a**) that differ in their intercalation topologies, the outer cytidine being either that on the 3' end or that on the 5'end of the stretch. Structures with different intercalation topologies may have comparable stabilities. Certain i-motif can also be formed in a dimer form (**Figure 1-11b**) of a DNA containing two cytidine stretches and an intermediate linker or by intramolecular folding of a single strand with four cytidine stretches (**Figure 1-11c**).[60] The intercalation and loop topologies are susceptible to the linker sequence.

Studies on certain oligonucleotides derived from fragments of natural sequences were carried out in several research groups in the past years to explore the possible biological or pharmaceutical relevance of i-motif.[58] One of the examples is the study of d(TAACCC), the sequence repeatedly occurred in human telomere. It was consequently demonstrated that this C-riched sequence existed in the form of i-motif when crystalised.[58] Furthermore, at each end of the i-motif core, one of the two TAA sequences loops could form an A-T pair that is stacked above the core.[56] Many efforts have also been made to study the intramolecular folding of DNA strands containing four cytidine stretches[61], the first high-definition structure of the sequence d(5m$CCT_3CCT_3ACCT_3CC$) was obtained from NMR studies in 1998.[63]

1.3.3 Biological Role of i-Motif Structure of DNA

A number of proteins that interact with the G-quadruplexes have been identified in the past,[32-40] in either unfolded or tetrad form. On the other hand, it is shown that i-motif may also play certain biological roles *in vivo*. For example, it was formed that some proteins could interact selectively with repeats of the telomeric cytosine-rich strand, possibly in the i-motif form.[64, 65] A protein of high molecular mass (160 kDa) that binds the cytosine-rich strand of the centromeric dodeca-satellite of Drosophila has also been characterized by Azorin and co-workers in 1999.[64] In addition, two human nuclear proteins, hnRNP K and the splicing factor ASF/SF2, were reported by Lacroix *et. al.* in 2000 [65] that bind to the telomeric cytosine-rich strand.

Chapter 2

Construction of i-Motif-Based DNA Machines

2.1 Background and Aims

2.1.1 Biomolecular Machines in Organisms

The concept of a macroscopic machine has been extended to the molecular level (molecular machine) in the past few decades.[66-69] Molecular machines can be defined as devices that conduct specific function through interactions of distinct molecular components.[66] Each molecular component performs a single action while the entire assembly performs a more complex function which can convert certain forms of energy to mechanical work.[67] Molecular machines are generally more efficient as compared with the macroscopic machinery. The machineries at the molecular level are generally fueled with chemical energy, electrical energy and photochemical energy for their mechanical actions. [67]

Nature creates its own set of molecular motors that have been working for millions of years inside the living systems. The bounden duty of these machines endowed by nature is to transform energy from one form to another in order to maintain cellular structures and functions. Most of these molecular machines are protein-based and rely on the energy-rich molecules (e.g. ATP) for their function. So far, two kinds of existing nature molecular machines have been well characterized. The first one is the F_0F_1-ATPase molecular motor (**Figure 2-1**). These motor proteins,

which are found in mitochondria, bacteria and chloroplasts, convert the energy stored in a transmembrane proton gradient to chemical energy (ATP).[68] The second type in this category is bacterial flagellar motors.[69] The movement of bacterial in most cases is driven by its flagella motors. The rotary motions of the flagella motors are commonly powered by protons flowing through cell membrane, a process in which chemical energy is transduced to mechanical motion.

These natural molecular machines have been developed and optimized during millions of years' evolutionary process. Even though they have adopted extremely complicated systems for their actions which are unlikely to be reproduced in the laboratory,[67] scientists have successfully made use of DNA molecules to build artificial molecular machines to mimic the natural molecular machinery. [88, 89]

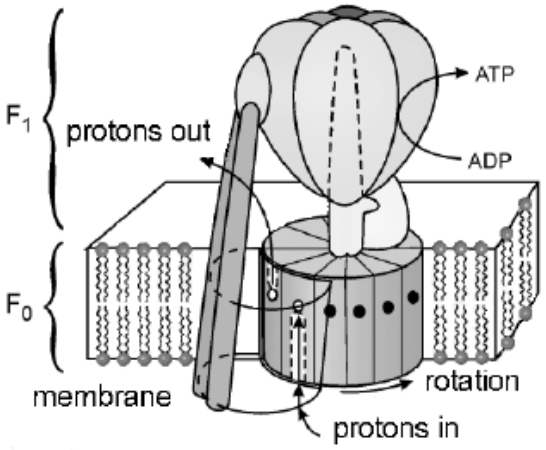


Figure 2-1. The F_0F_1-ATPase molecular motor

2.1.2 DNA-based Artificial Molecular Machines

Inspired by the wonder of biomolecular machinery found in nature, much attention has been paid to the development of artificial molecule machine which can fulfill certain mechanical functions.[70-73] Because of some unique characteristics of DNA, this type of biomacromolecule has been considered as an ideal building block

in molecular nanotechnology.[70] For example, one of the distinctive properties of DNA is the elastic properties of double strand DNA (dsDNA) which leads to desired bending rigidity and twisting rigidity of the biomacromolecule.[70] In addition, DNA can be designed to self-assemble in a preferable fashion.[71-73] Moreover, since some of the DNA structures possess conformational isomers, control of transformation between isomers are attractive for the fabrication of switching devices. Various DNA sequences have therefore been utilized over the years for the construction of molecular wires, molecular grids, and other nanoscale molecular objects.

One of the most innovative DNA-based molecular switches was developed by Seeman and co-workers in 1999[74]. This DNA switch is comprised of two rigid DNA double-crossover (DX) motifs (**Figure 2-2**). The reversible transformation between B-form (right-handed) and Z-form (left-handed) of DNA can be triggered by additional of metal ions. The main strand of this molecule switch contains two short $d(CG)_{10}$ domains, which are designed to form intramolecularly a duplex structure in the middle of the DNA motif. Two fluorescent probes are incorporated site-specifically into the DNA motif at certain positions. The duplex section is transformed from B to Z upon addition of $Co(NH3)_6^{3+}$ which is accompanied by twisting of the DNA motif at duplex domain.

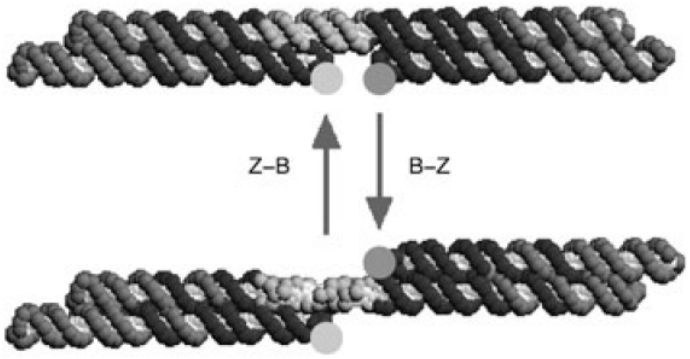


Figure 2-2. DNA-based twisting molecular switch

In addition, Yurke and co-workers[75] designed certain artificial molecular machines based on intermolecular DNA hybridization. This special device mimics the function of tweezers at the molecular level through using three strands of DNA (DNA tweezers, **Figure 2-3**). In this system, strand A were hybridized to each end of strands B and C to form two arm-like double-stranded structures (**Figure 2-3**).[75] In its open conformation, fluorophore and quencher are separated apart. Addition of the DNA strand F, which is complementary to the single stranded form of DNA could result in a closure of the "open" tweezers to generate a compact configuration (state 5 in **Figure 2-3**). This effect is reversed by the addition of strand F', which is fully complementary to F.

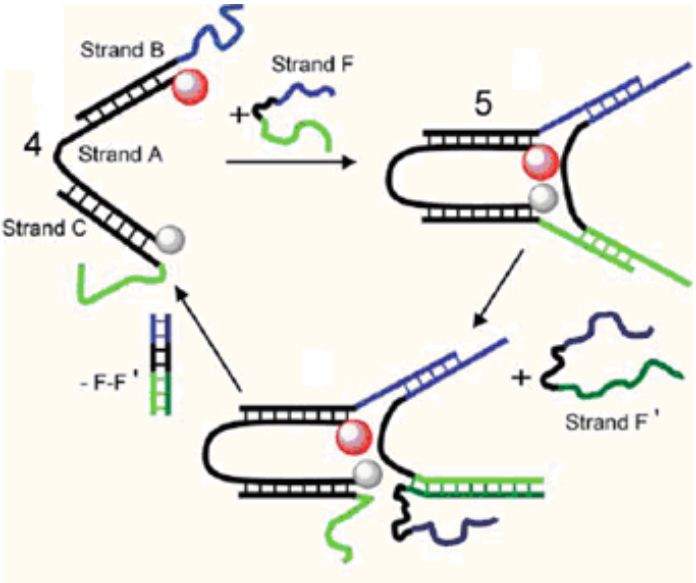


Figure 2-3. DNA tweezers [75]

Using similar principles, Simmel and Yurke[76] developed a DNA nanoactuator which can switch from a relaxed, circular form to a stretched conformation. Additional example of DNA devices designed in the recent years is called "DNA scissors".[77] Mitchell and Yurke demonstrated that two sets of tweezer structures could join at their hinges with short carbon linkers under proper conditions. The motion of

one set of tweezers could be transduced to the other end of the DNA molecules, resulting in a scissor-like movement.[77]

Since 2004, several research groups have developed some new molecular motors that can direct DNA to walk along designed routes, namely DNA walkers.[78] These DNA walkers move along single-stranded nucleic acid tracks and are fixed to the track throughout the entire operation. The actions of some of the DNA walkers are illustrated in **Figure 2-4**. [78]

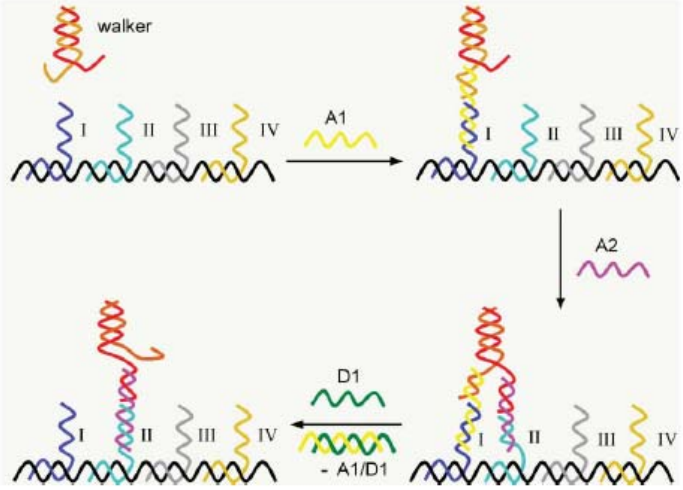


Figure 2-4. DNA walkers [78]

2.1.3 Quadruplex DNA-based Molecular Machines

Structural competitions between G-quadruplex, i-motif and duplexes have been investigated by several research groups. For example, Phan & Mergny[79] examined a human telomeric fragment d[AGGG(TTAGGG)$_3$] / [d(CCCTAA)$_3$CCCT] under a variety of experiment conditions and studied the conversion of the regular double-helix structure into the intramolecular G-quadruplex and i-motif. They demonstrated that DNA predominantly exists in the double-helix form under near-physiological condition while at lower pH or higher temperatures, duplex structure could dissociate and G-quadruplex and i-motif will form. Furthermore, Li *et. al.*[80]

observed that under certain circumstances, formation of quadruplex and duplex of DNA depended on thermal stabilities of the corresponding DNA assemblies. Moreover, Risitano & Fox[81] observed that stability of the quadruplex does not increase with the increase of length of G-tract and linking bases. All these investigations imply that G-quadruplex DNA could be more favorable under specific conditions.

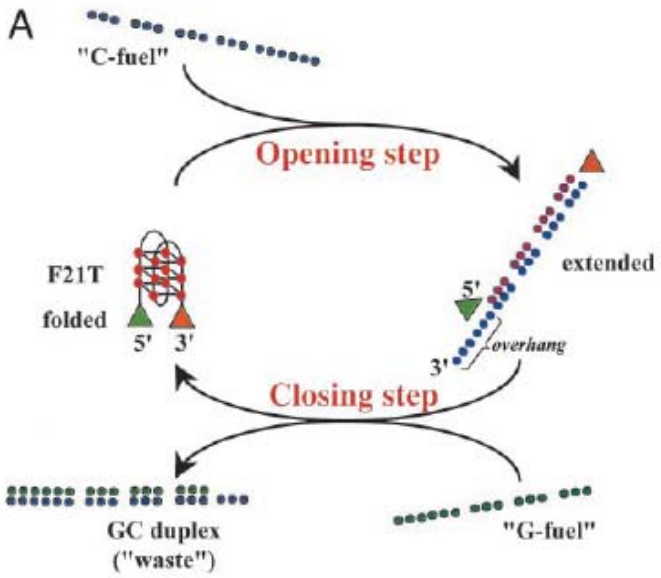


Figure 2-5. A quadruplex-duplex exchange nanomachine [82]

Taking advantage of unique structural properties of i-motif and G-quadruplex forms of DNA, scientist have developed different types of nanodevices using these tetraplex forms of DNA since 1990s.[87-89] For example, Alberti & Mergny[82] introduced "C-fuel" and "G-fuel" into a nanodevice and demonstrated the manipulation of an extension-contraction cyclic movement. This simple device is composed of a single 21-mer oligonucleotide and relies on the duplex / quadruplex equilibrium for its action. The single strand is initially folded into a qudruplex structure and subsequent addition of a complementary C-strand will force it to form a duplex assembly. After addition of G-fuel which is complementary to C-fuel, the corresponding G-rich strand

will be released and further self-assemble into a quaduplex structure (**Figure 2-5**).[82]

Simmel and co-workers extended this approach to biological systems and used the structure of aptamer in their construction of a nanomachine, which can bind and release a single protein molecule.[83] The aptamer sequence used in Simmel's studies could fold into a two layers G-quartets in the presence of potassium ion and bind strongly to a human blood-clotting protein, thrombin. As shown in **Figure 2-6**, upon addition of single-stranded DNA that is complementary to a portion of the aptamer sequence, a duplex structure is formed, which leads to the release of the corresponding protein. The operation of this machine can be monitored using FRET. In this aptamer device, formation and dissociation of G-quadruplex entities were involved and further utilized.[83]

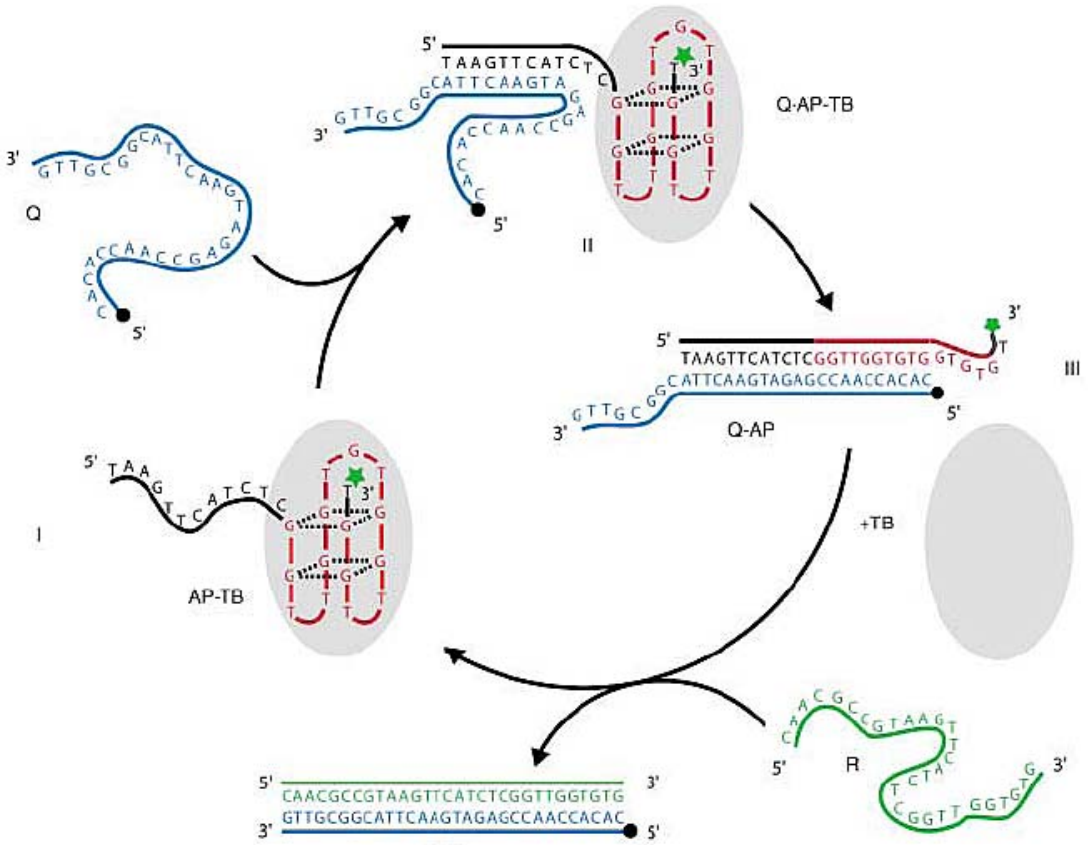


Figure 2-6. A switchable aptamer device [83]

In addition, Bourdoncle *et. al.* [84] built the quadruplex-based molecular beacons

using G-quadruplex as stem in which a central loop is composed of 12 to 21 bases. The central loop hybridizes and forms a duplex structure which helps opening process of the quadruplex stem. The quadruplex-based molecular beacon (G4-MB) can be formed during the equilibrium process. Detailed thermodynamic and kinetic investigations in these new systems were also carried out by these researchers.[84]

Besides the widely used strand displacement strategy, variation of environmental medium was utilized in the design of the quadruplex-DNA-based nanomachines. For example, Liedl & Simmel[85] constructed a chemical oscillator via controlling the formation and dissociation of i-motif structures. This specific oscillatory reaction could produce pH variations in the range between pH 5 and 7. Association and dissociation of i-motif structure was accordingly regulated via variation of pH within this pH range. The cytosine-rich DNA sequence is transformed between a random coil conformation and a folded i-motif structure under this condition.

In addition, through manipulating molecular devices via chemical reactions reported by Simmel *et. al.*, physical regulation could also be adopted in the design of molecular devices. S. Balasubramanian and co-workers[86] used protons to fuel a nanomachine which can be controlled between an i-motif conformation and an extended double-stranded structure (**Figure 2-7**). At low pH values, self-assemble of i-motif structure on the basis of $C\text{-}C^{+}$ interaction are involved; At higher pH values, cytosine is deprotonated and X strand in **Figure 2-7** can hybridize with the complementary strand Y to form a duplex structure (open state).

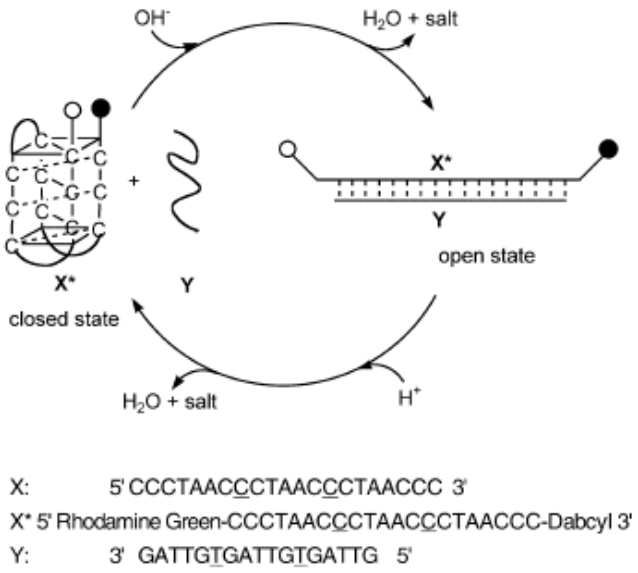


Figure 2-7. A Proton-Fuelled DNA Nanomachine[86]

Even though various artificial DNA machines have been developed in the past [87-89] including those designed on the basis of G-quadruplex, application of i-motif structure in the design of molecular devices has not yet been well explored. Inspired by the natural beauty of molecular machines adopted by organism and with the intention of exploiting the possibility of designing new DNA machines which can perform well-defined actions (e.g. Watson-Crick breakage), an i-motif-based DNA machine has been designed and constructed during my graduate studies that can take in chemical energy associated with acid-base reactions and further convert it to a new type of mechanical work. Design, synthesis and operation of this i-motif-based DNA machine are discussed in the following sections of this chapter.

2.2 Our Strategies in Design of i-Motif-Based DNA Machines

Let us imagine that there is a two-way electric winch that is hooked to the ends of a bow (**Figure 2-8**). The electric winch takes in its cables gradually; the two ends of the bow will be forced to move toward each other in a steady manner. If the

main body of the bow is tied up simultaneously to a still wall with breakable ropes, these fragile linkages will have to endure a force originated from the arm movement of the bow. Over this course of action, the electric winch converts electrical energy into mechanical work that leads ultimately to the fractures of the fragile ropes between the bow and wall.

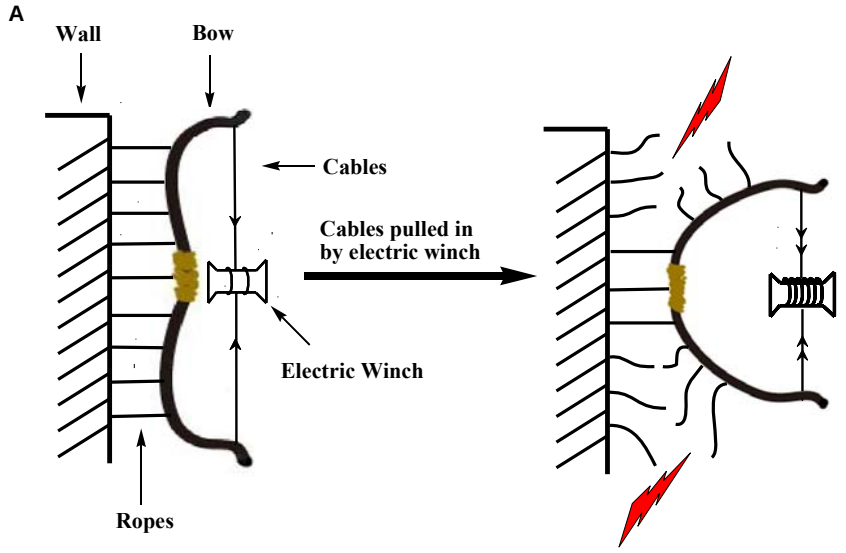


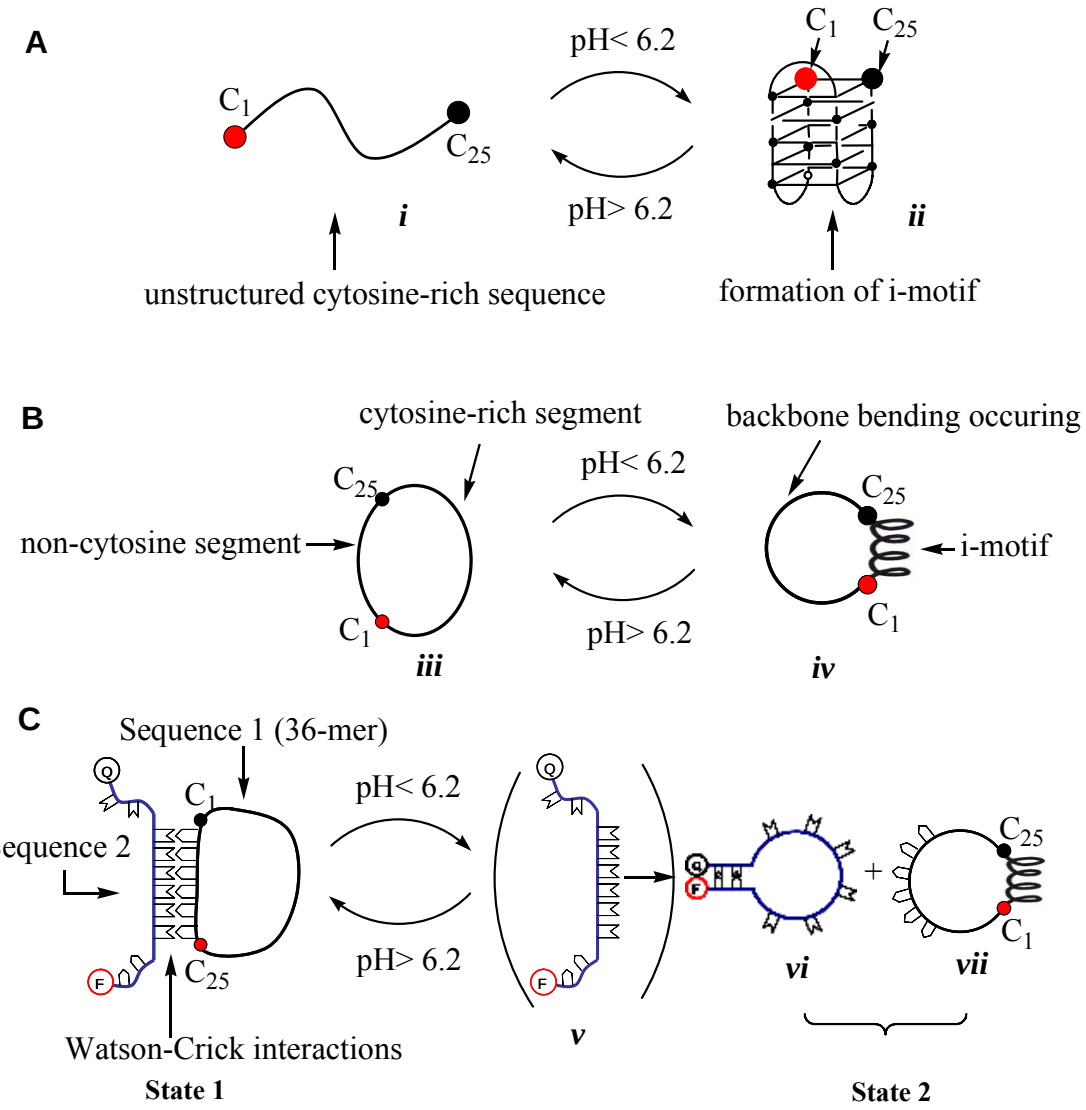
Figure 2-8. Illustration of our designed DNA-based molecular machine

Similar to the connecting fashion between the wall and bow on the macroscopic scale to a certain extent, duplex structures of two complementary DNA strands are held together via Watson-Crick interaction, which are inseparable unless certain forms of external energy are imposed on them. During replication, transcription and other genetic processes in vivo, dissociation of these duplex forms of DNA are executed by helicase,[90] a protein motor that takes up chemical energy derived from nucleotide hydrolysis.[91] Our approach to break down Watson-Crick interaction in this study is to covalently link two termini of one of the two complementary oligonucleotide strands to a cytosine-rich sequence. Under a neutral condition, the two ends of a cytosine-rich oligonucleotide will position themselves randomly and might not be able to get close to each other readily due to the rigidity of

DNA backbone.[92] Formation of a structural entity of i-motif under an acidic condition, however, will force the two ends of this cytosine-rich sequence to move towards each other.[92] It was envisioned during our early investigations that in the process of formation and disassociation of i-motif, a cytosine-rich sequence could act in the same fashion as an electric winch does on the macroscopic scale (**Figure 2-8**).

A duplex structure-containing complex was accordingly designed during our investigations that was constituted of a circular oligonucleotide of 36-mer (<C_1CCCT_5TTCCC$_{10}$CTTTC$_{15}$CCCTT_{20}TCCCC$_{25}$AAAAT_{30}TAAAA$_{35}$A>, **Sequence 1**) and a linear oligonucleotide of 19-mer tagged with BODIPY and DABCYL at its 5' and 3' ends (5' BODIPY-ATATTTTTTAATTTTATAT-DABCYL 3', **Sequence 2**). This circular sequence (**Sequence 1**) contained both a cytosine-rich stretch (25 nucleotides in length) and a non-cytosine-containing tract (11 nucleotides in length) respectively, the former possessing an innate ability to form a structural assembly of i-motif, and the latter being complementary partially to **Sequence 2**. It was our anticipation that under a neutral condition, **Sequence 1** and **Sequence 2** would be held together by Watson-Crick interaction in a tight fashion (State 1 in **Figure 2-9C**) via the two complementary segments that they contain. As pH decreased, the cytosine-rich stretch of **Sequence 1** would form a structural entity of i-motif,[92] which would consequently drive Nucleotide A_{26} and Nucleotide A_{36} of **Sequence 1** to move close to each other (State 2 in **Figure 2-9C**). This forced movement of A_{26} and A_{36} could, in theory, cause sever backbone bending on the non-cytosine segment (-A_{26}AAAT_{30}TAAAA$_{35}$A-) of this circular oligonucleotide. Since there was no additional physical force to compel the backbone of **Sequence 2** to bend along with the non-cytosine segment of **Sequence 1**, duplex structure between these two sequences would be disintegrated (**Figure 2-9C**) as a consequence. During this

course of action, chemical energy associated with acid-base reactions would be converted into physical work that leads to the breakdown of Watson-Crick interaction.



Sequence 1 (circular 36-mer):

<C_1CCCT$_5$TTCCC$_{10}$CTTTC$_{15}$CCCTT$_{20}$TCCCC$_{25}$AAAAT$_{30}$TAAAA$_{35}$A>

Sequence 2 (linear 19-mer tagged with a fluorophore and its quencher):

5' Bodipy -ATATTTTTTAATTTTATAT- Dabcyl 3'

Figure 2-9. Schematic representation of our strategy for designing a new energy-converting DNA machine capable of breaking down Watson-Crick interaction.

(A) Formation and dissociation of structural entity of i-motif as the pH value of its environment varies.
(B) Transformation of chemical energy coupled with acid-base reaction to elastic potential energy within a circularly-backboned DNA.
(C) Transformation of chemical energy coupled with acid-base reaction to electrical potential energy and breakdown of Watson-Crick interaction.

As pH changes from acidic back to neutral, the elastic potential energy stored in the form of backbone bending will be librated to its environment as the structural entity of i-motif falls apart. Under neutral condition, **Sequence 1** and **Sequence 2** will form an 11-mer duplex structure via their complementary segments, an entity sustained by hydrogen bonding (electrostatic attraction) (**State 1** in **Figure 2-9C**). As pH decreases, a high degree of curvature on the non-cytosine segment will be forced to occur by the formation of structural entity of i-motif. Owing to the absence of extra force to drive **Sequence 2** to curve with its complementary segment, the corresponding Watson-Crick base pairs will be detected, leading to an electrical potential energy increase of the system in the form of dissociated dipolar groups preserved in the single stranded oligonucleotides (***vi*** and ***vii*** in **Figure 2-9C**). At the same time as electrostatic interactions are broken down, the resultant single stranded form of **Sequence 2** (***v***) will adopt a stem-and-loop conformation (***vii***) in which fluorescence signal of Bodipy is quenched to a high extent. When the elastic potential energy stored in ***vii*** is allowed to liberate next by increasing the pH of the system, the electrical potential energy stored in the single stranded form of this oligonucleotide will be discharged by regenerating electrostatic attraction between the two complementary segments from **Sequence 1** and **Sequence 2** respectively. At the same time as a duplex structure is reestablished, fluorescence signal of Bodipy is being restored.

2.3 Synthesis of Our Newly Designed i-Motif-Based DNA Machines

Sequence 1 was accordingly synthesized through the enzymatic ligation reaction of two termini of a linear 36-mer oligonucleotide (5’ TAAAAACCCCTTTCCCCTTTCCCCTTTCCCCAAAAT 3’, **Sequence 3**) catalyzed

by T4 DNA ligase in the presence of a template sequence (**Sequence 4**, GGTTTTTAATTTTGG). The reaction mixture containing 1 μM **Sequence 3**, 3 μM **Sequence 4**, 1 x T4 DNA ligase buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP, 25 μg/ml BSA, pH 7.8), 30 mM NaCl and 100 U T4 DNA ligase were incubated at 20 °C for different periods of time. Radioactive strand of **Sequence 3** was prepared by phosphorylation of the synthetic sequence at 5' end with T4 polynucleotide kinase and [γ-^{32}P] ATP. The phosphate transfer reaction between [γ-^{32}P] ATP and the 5' hydroxyl-ended **Sequence 3** was catalyzed by the kinase. (Refer to "**6.2.1** 5' End labeling of DNA" for detailed procedures.) The ratio of the linear precursor and the template sequence is 1:3. The circularization was performed at 20 °C, under which the duplex structure could be sustained. It takes about 1 hour to complete the circularization reaction. As shown in **Figure 2-10A**, a clear band with lower mobility shift than the linear precursor appeared on the gel which was proved to be the circular product in later studies.

The circular nature of this newly formed oligonucleotide in its backbone was subsequently confirmed by examining its resistance to hydrolysis by exonuclease I. The reaction mixture containing 1 x exonuclease I buffer (67 mM Glycine-KOH, 6.7 mM $MgCl_2$, 10 mM 2-mercaptoethanol, pH 9.5), oligonucleotides (**Sequence 1** or **Sequence 3**) and 80 U exonuclease I was incubated at 37 °C for 60 min. As shown in **Figure 2-10B**, the newly formed product was completely resistant to the enzyme hydrolysis. As a control, linear precursor was degraded as small fragments.

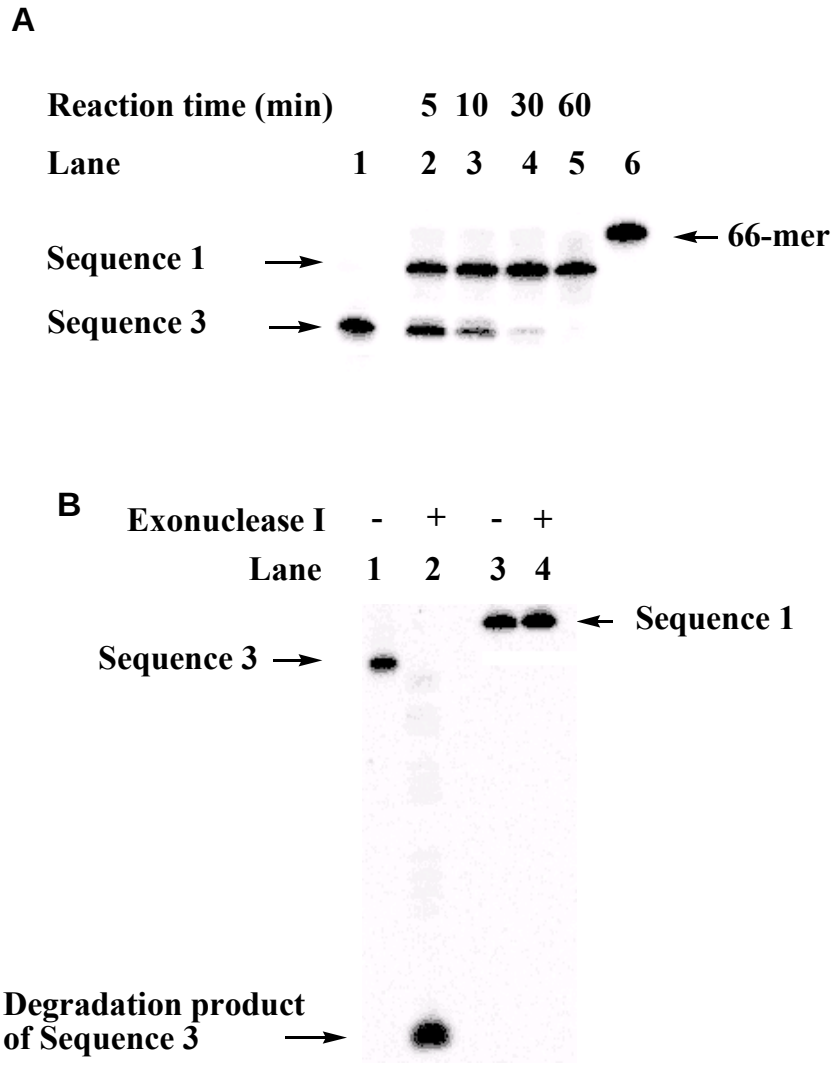


Figure 2-10. Polyacrylamide gel electrophoretic analysis of oligonucleotides as components of the artificial DNA machines designed in our study.

(A) Synthesis of **Sequence 1**. Lane 2: 5 min, lane 3: 10 min, lane 4: 30 min, lane 5: 60 min. (B) Confirmation of circularity of **Sequence 2** in its backbone. Lane 1: **Sequence 3** in the absence of exonuclease I (as control), lane 2: **Sequence 3** in the presence of exonuclease I, lane 3: **Sequence 1** in the absence of exonuclease I (as control), lane 4: **Sequence 1** in the presence of exonuclease I.

2.4 Operation of Our i-Motif-Based DNA Machines

2.4.1 First Half and Second Half of Operating Cycle

It is known that BODIPY dyes possess: [93]

a) High extinction coefficients (>60,000 cm^{-1} M^{-1});

b) High fluorescence quantum yields (often approaching 1.0, even in water);

c) Fluorescence is pH-insensitive;[94]

d) No ionic charge;

e) Narrow emission bandwidths; and

f) Greater photostability than fluorescein in some environments.

In addition, it was demonstrated by Vos de Wael, *et. al.*[95] that the BODIPY 493/503 fluorophores are brilliantly fluorescent in a wide range of solvents. A BODIPY 493/503 was consequently selected during our investigation which was linked to 5' end of particularly designed oligonucleotides (**Figure 2-11**). This 5' end labeled Bodipy 493/503 has a maximum of fluorescence emission of 515 nm when it is excited at 480 nm[86] and is insensitive to pH variation from 2 to 8.5.[93] On the other hand, DABCYL was linked to 3' end of the oligonucleotide, which serves as a quencher of the corresponding fluorophore.

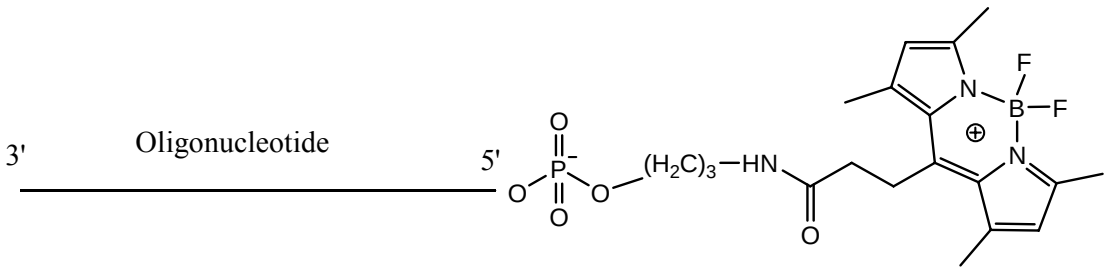


Figure 2-11. Structure of Bodipy 493/503 modification on 5' end of oligonucleotide

A sample containing **Sequence 1** and BODIPY-tagged **Sequence 2** in an equal ratio was accordingly examined next at 20 °C using fluorescence spectroscopy. As shown in **Figure 2-12A**, the emission intensity of this sample at 515 nm increased from 83 a.u. to 503 a.u. with the increase of pH from ~5.5 to ~7.5. These observations were consistent with the suggestion that **Sequence 2** existed in a stem-and-loop conformation at ~pH 5.5 (Structure 1 in **Figure 2-9C**) and was further transformed into a stem-open duplex structure with **Sequence 1** at ~pH 7.5 (State 2 in

Figure 2-9C). Subsequent variation of pH of the same system from ~7.5 to ~5.5 led to a decrease of emission intensity from 493 a.u. to 87 a.u. (**Figure 2-12B**). This dramatic reduction of emission intensity indicates that the stem-open form of **Sequence 2** was further transformed back to its stem-and-loop conformation (from State 1 to State 2 in **Figure 2-9C**) under an acidic condition.[97] These conformational alterations of **Sequence 2** implied that a breakdown of Watson-Crick interaction between **Sequence 1** and **Sequence 2** took place during the course of pH variation.

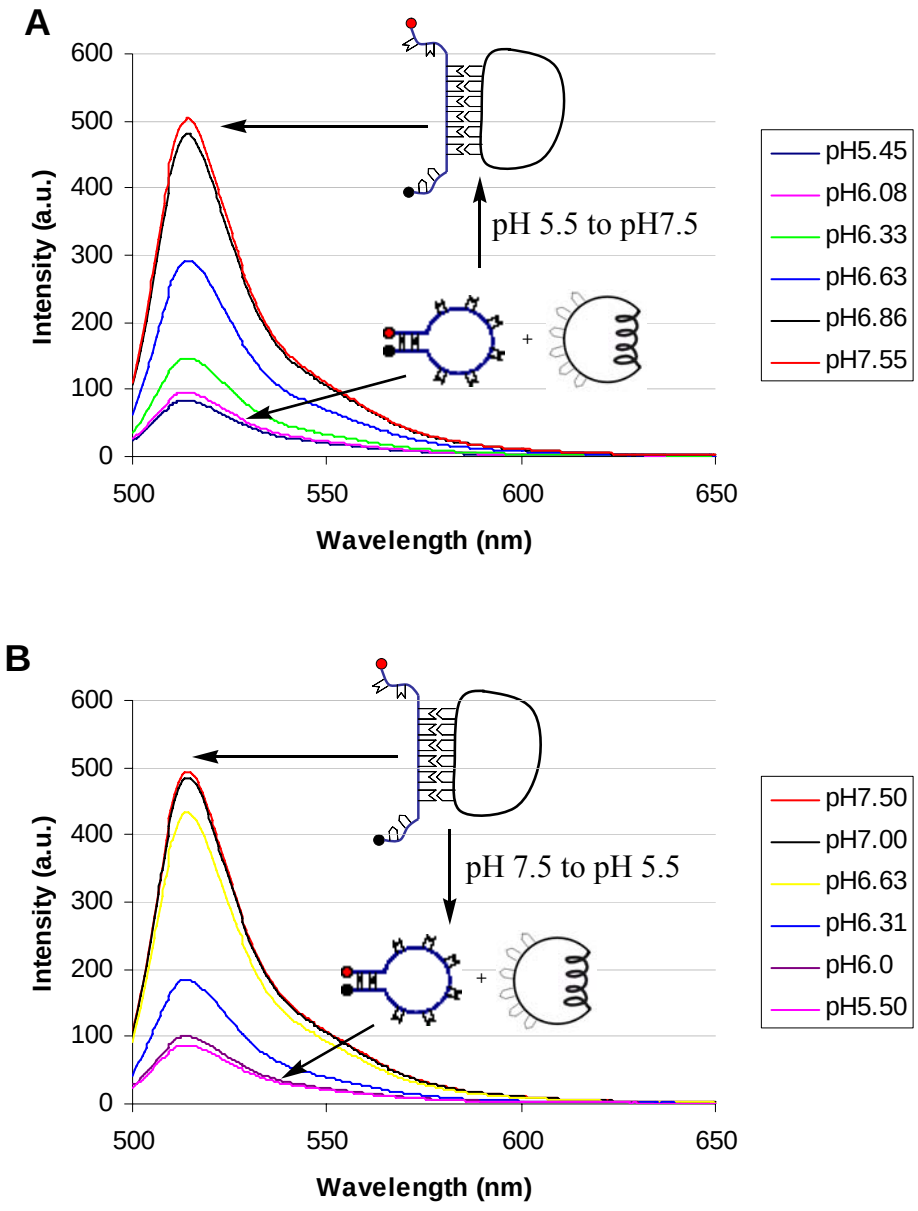


Figure 2-12. Fluorescence Spectroscopic analysis of formation and disintegration of duplex structure associated with the artificial devices designed in the current studies.

(A) Emission spectra of **Sequence 1** and **Sequence 2** as pH increases. Fluorescence spectra of a sample containing 1 μM **Sequence 1**, 1 μM **Sequence 2**, 10 mM MES/HEPES (pH 7.5) and 1 M NaCl was recorded at 20 °C with excitation at 480 nm. Variation of pH was carried out by addition of 1 M NaOH to the system in a stepwise fashion.

(B) Emission spectra of **Sequence 1** and **Sequence 2** as pH decreases. Fluorescence spectra of a sample containing 1 μM **Sequence 1**, 1 μM **Sequence 2**, 10 mM MES/HEPES (pH 5.5) and 1 M NaCl was recorded at 20 °C with excitation at 480 nm. Variation of pH was carried out by addition of 1 M HCl to the system in a stepwise fashion.

With the aim of confirming that the changes of emission intensity shown in **Figure 2-12A** and **Figure 2-12B** were indeed correlated with the formation of i-motif, a sample consisting of **Sequence 2** and its complementary 11-mer of non-cytosine-containing oligonucleotide (5' AAAATTAAAAA 3', **Sequence 5**) was examined under different pH conditions using fluorescence spectroscopy. In theory, **Sequence 2** and **Sequence 5** would assemble together to form a duplex structure under a neutral condition while such a structural entity would be maintained in an acidic environment. As seen in **Figure 2-13b**, there was little alternation of emission intensity occurred when the pH of a sample containing **Sequence 2** and **Sequence 5** decreased from ~7.5 to ~5.5 as well as increased from ~5.5 to ~7.5. These observations demonstrate that a cytosine-rich segment is an essential component of the DNA machine designed in the studies. In addition, the absence of dramatic change of emission intensity with pH variation (**Figure 2-13b**) rules out the possibility that the alterations of emission intensity in **Figure 2-12A** and **Figure 2-12B** are caused by the protonation of the thymines and/or adenines within **Sequence 1** and **Sequence 2**.

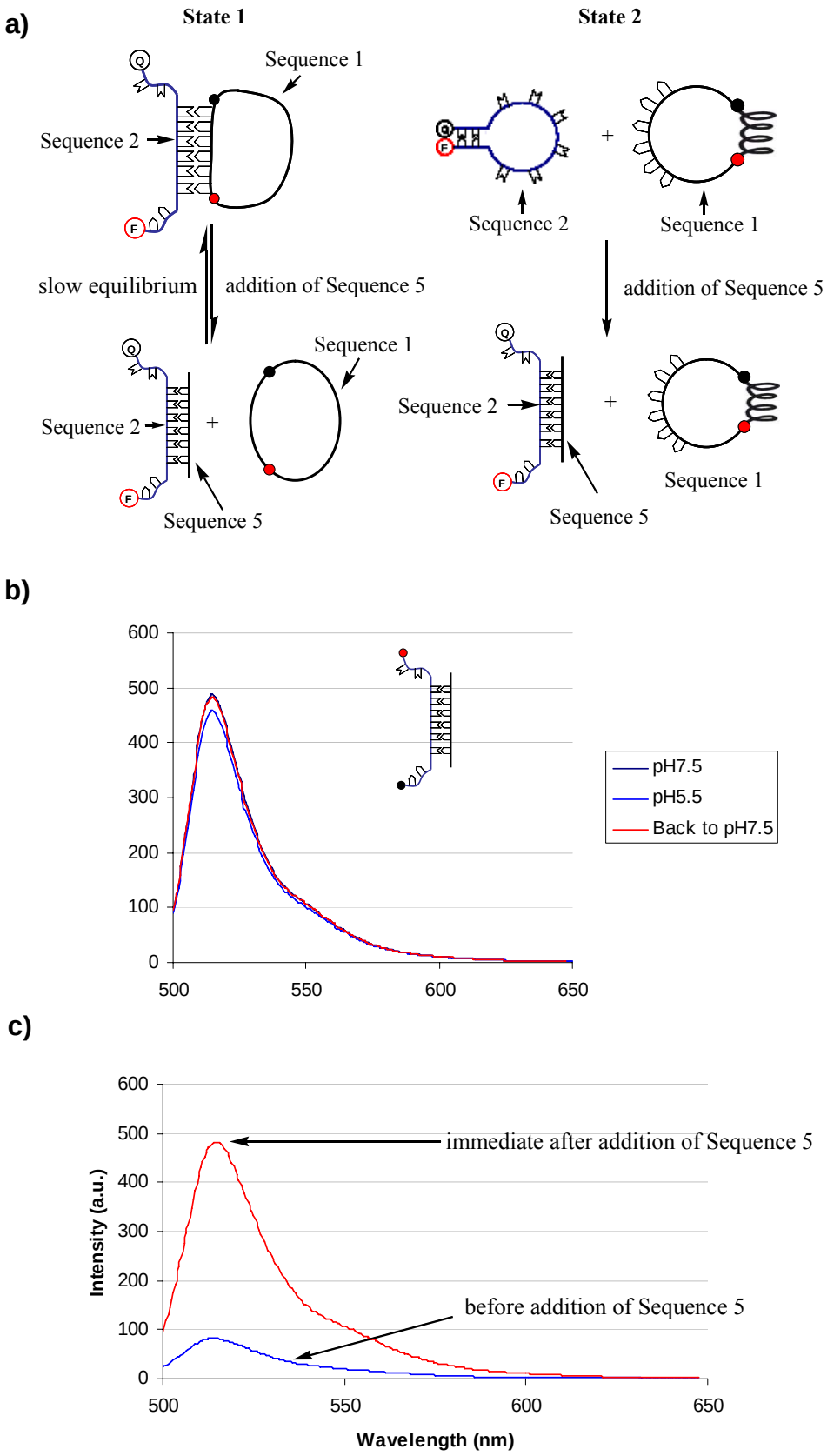


Figure 2-13. Analysis of dissociation and formation of duplex structure correlated with the artificial machines using fluorescence spectroscopy.

a) Illustration of correlation among **Sequence 1**, **Sequence 2** and **Sequence 5** under different conditions.

b) Emission spectra of **Sequence 2** and **Sequence 5**. A mixture containing 1 μM **Sequence 2**, 1 μM **Sequence 5**, 10 mM MES/HEPES and 1 M NaCl at pH 7.5 was prepared and its emission spectrum was recorded at 20 °C with excitation at 480 nm (Black). 1 M HCl was added next to the above solution until pH reached ~5.5 (Blue) followed by the adjustment of pH of the new mixture to pH ~7.5 through adding 1 M NaOH (Red).

c) Emission spectra of **Sequence 1**, **Sequence 2** and **Sequence 5** in different combinations. A mixture containing 1 μM **Sequence 1**, 1 μM **Sequence 2**, 10 mM MES/HEPES (pH5.5) and 1 M NaCl was prepared and its emission spectrum was recorded at 20 °C with excitation at 480nm (Blue). An equal amount of **Sequence 5** to that of **Sequence 1** was added next to the above solution and an emission spectrum of the new solution was further recorded under the same condition (Red).

Mse I is a restriction endonuclease that recognizes TTAA-containing double helices of DNA and makes its incisions within this duplex segment[98] (detailed information refer to **Chapter 6.1.2**). To further prove that a duplex structure between **Sequence 1** and **Sequence 2** did indeed exist under neutral condition (State 1 in **Figure 2-9C**), a DNA cleavage test with *Mse I* was carried out during our investigations. As shown in **Figure 2-14**, the circular backbone of ^{32}P-labeled **Sequence 1** was converted into its linear forms when *Mse I* was added to a mixture containing both **Sequence 1** and **Sequence 2** at pH 7. This observed strand scission is the indication of presence of Watson-Crick interaction in the designated system as designed (State 1 in **Figure 2-9C**). In addition, with the purpose of further verifying that **Sequence 1** was truly separated from **Sequence 2** under acidic conditions (State 2 in **Figure 2-9C**), **Sequence 5** (5' AAAATTAAAAA 3', complementary partially to **Sequence 2**) was added to a sample containing preincubated **Sequence 1** and **Sequence 2** at pH 5.5. As shown in **Figure 2-13c**, this addition caused an instant and dramatic increase of fluorescence intensity of the sample, which suggests that a stem-and-loop conformation of **Sequence 2** did preexist in the system as originally

intended.

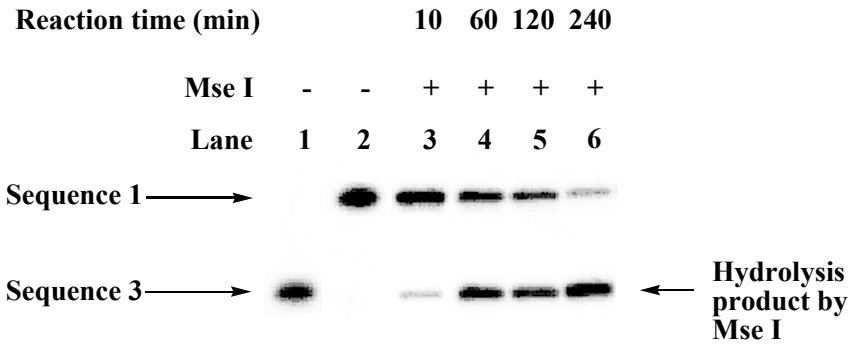


Figure 2-14. Confirmation of presence of duplex structure between **Sequence 1** and **Sequence 2** (State 1 in **Figure 2-9C**) at pH > 6.2. A solution containing 1 μM **Sequence 2**, 1 μM ^{32}P-labled **Sequence 1**, 1 x Mse I buffer (10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl2 1 mM Dithiothreitol, pH 7.9), 100 mM NaCl and 60 U Mse I was incubated at 20 °C for different time periods. Lane 3: 10 min, lane 4: 60 min, lane 5: 120 min, lane 6: 240 min.

2.4.2 Cyclic Operation of i-Motif-based DNA Machine

Since assembly and disassembly of i-motif and duplex structures of DNA were involved in the functioning process of the DNA machine designed in this study, the time scales needed to complete each half of the cycle of the operation were measured next. The obtained response times of this DNA machine to both pH decrease and increase were around 10 seconds (**Figure 2-15**, **A** and **B**), which matched up with those of similar operations reported in literature.[86, 99] Given that a relatively short time period was needed for each half of the cycle of the operation, this DNA machine was further run in a periodic fashion. For the first half of the cycle, 1 M NaOH was added to a solution containing 1 μM **Sequence 1**, 1 μM **Sequence 2**, 10 mM of MES/HEPES (pH 7.5) and 1 M NaCl in stepwise fashion until pH reached ~5.5. 1 M HCl was added next to the above solution until pH reached ~ 7.5 to complete the second half of the cycle. Emission intensity at 515 nm was recorded at 20 °C in every 1 min with excitation at 480 nm. As shown in **Figure 2-15C**, decrease in fluorescence

emission of the system is insignificant over five cycles, indicating that the accumulation of waste products generated from acid-base reactions has little effect on the energy-converting process in a limited number of cyclic operations.

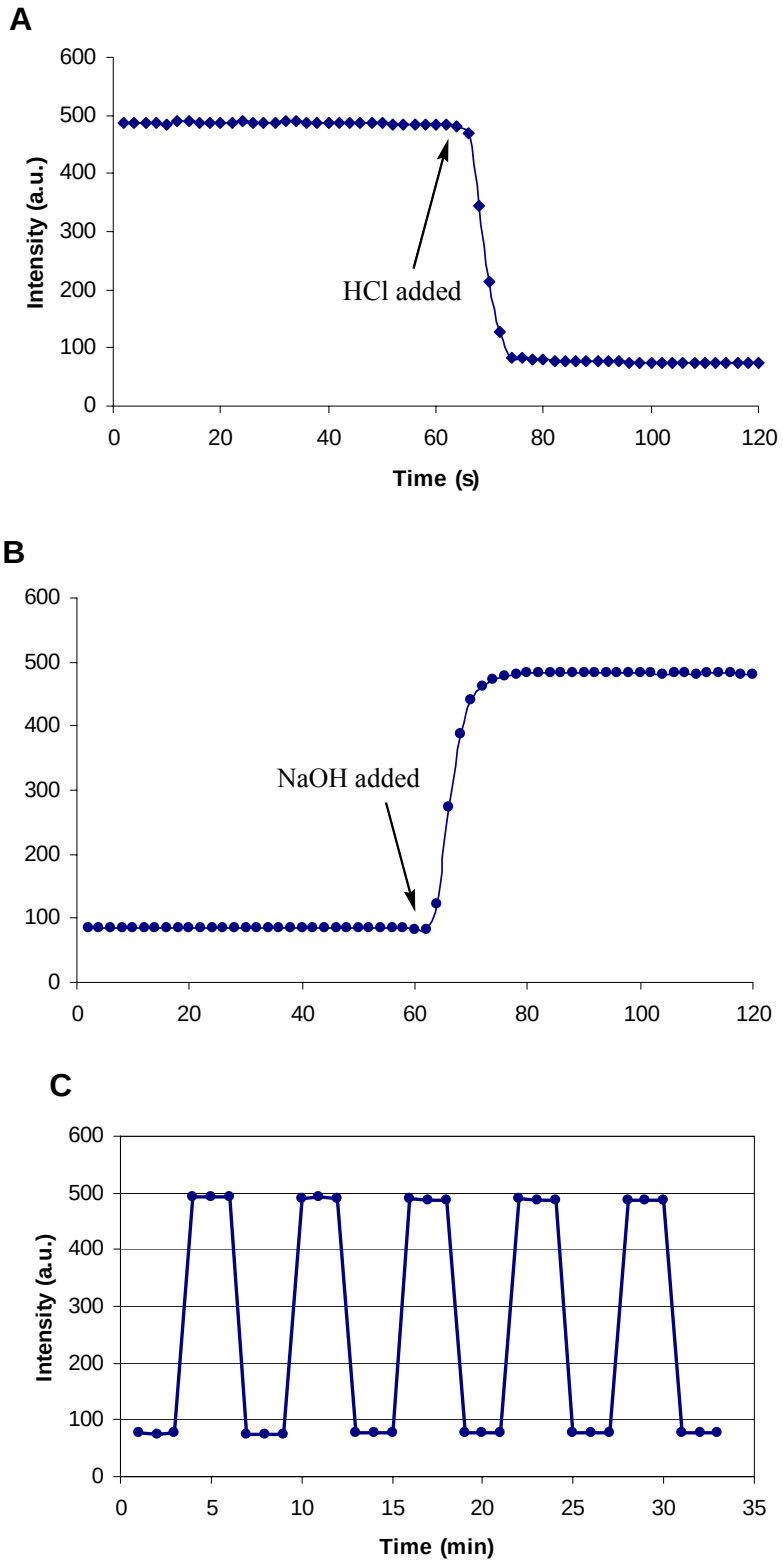


Figure 2-15. Examination of operability of artificial DNA machines using fluorescence spectroscopy.

(A) Changes of emission intensity in the first half of the cycle as pH decreases from 7.5 to 5.5. 1 M HCl was added to a solution containing 1 μM of **Sequence 1**, 1 μM of **Sequence 2**, 10 mM MES/HEPES (pH 7.5) and 1 M NaCl in a stepwise fashion until pH reached ~5.5. Emission intensity at 515 nm was recorded at 20 °C in every 2 s with excitation at 480 nm.

(B) Changes of emission intensity in the second half of the cycle as pH increases from 5.5 to 7.5. 1 M NaOH was added to a solution containing 1 μM of **Sequence 1**, 1 μM of **Sequence 2**, 10 mM of MES/HEPES (pH 7.5) and 1 M NaCl in stepwise fashion until pH reached ~7.5. Emission intensity at 515 nm was recorded at 20 °C in every 2 s with excitation at 480 nm.

(C) Changes of emission intensity as the designed DNA machines were operated for five cycles.

2.4.3 Calculation of Mechanical Energy Released by Our i-Motif-Based DNA Machine

Thermal stability of the 11 base pairs (5' AAAATTAAAAA 3'/5' TTTTAATTTTT 3') between circular **Sequence 1** and molecular beacon was also examined by using UV/VIS Spectroscopy. The observed melting point of this duplex entity under the conditions adopted in the current work is 37.8 °C (**Figure 2-16**), which is compatible with the data (35.0 °C, 1 μM Oligo Concentration and 1M NaCl) calculated on the basis of empirical estimation (refer to **Chapter 6.2.9**). All the experiments were performed at 20 °C to avoid the thermal denaturing of the complex.

In addition, according to theoretical calculations,[100] the free energy change for the formation of this duplex structure from its single-stranded forms is -12.46 kcal mol^{-1} (**Table 2-1**). This datum implied that the *work* done in the Watson-Crick interaction-breaking process designed in the current studies is 12.46 kcal mol^{-1}.

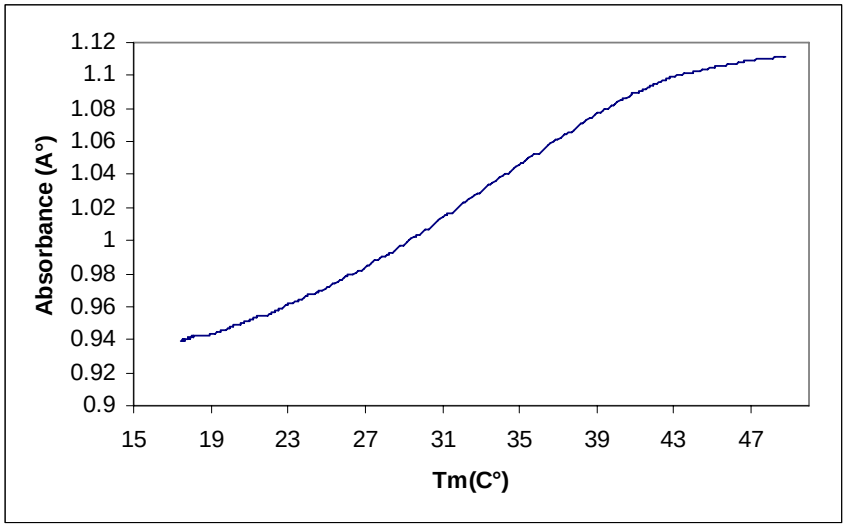


Figure 2-16. UV melting curve of the 11 base pairs duplex entity

Table 2-1 Calculations of the free energy changed during the formation of duplex structure from its single-stranded form.

Ct[Oligo] (mol/L)	Tm (°C)	Tm (K)	RTmln(Ct/4)	dA	ΔS (cal/mol.K)	ΔH (kcal/mol)	ΔG(20°C) (kcal/mol)
1.00E-07	29.3	302.5	-43994.82				
2.00E-07	31	304.2	-42490.19	1504.64	211.74	74.57	-12.49
5.00E-07	33.3	306.5	-40478.08	2012.11	209.29	73.82	-12.47
1.00E-06	35	308.2	-38927.66	1550.41	218.18	76.55	-12.59
1.50E-06	36	309.2	-38012.34	915.33	218.98	76.79	-12.60
2.00E-06	36.8	310	-37369.72	642.61	192.17	68.50	-12.17
							∑-12.46

2.5 Conclusions

When a DNA molecule responds to the changes of its environment, its conformation often adjusts. In order to link the conformation changes of DNA with desirable forms of energy transformation, a new approach of utilization of backbone-circularized DNA is introduced into the design of DNA machines during our investigation. It is consequently demonstrated in the current studies that chemical energy coupled with acid-base reaction can be transduced into elastic potential energy as well as electrical potential energy being preserved in the macromolecules involved. As a manifestation of these energy-converting processes, breakdown of Watson-Crick interaction is forced to occur under isothermal condition.

The outcomes of our studies not only illustrate the potentiality and versatility of DNA molecules as building blocks for fabricating sophisticated supramolecular architectures, but also further ascertain the manipulability of single molecules of DNA in terms of force, motion and energy transformations. It is our expectation that development of the new energy-converting DNA-based machine could have significant implications in the future design of new DNA architectures to carry out more intricate energy-converting actions.

Chapter 3

Search and Confirmation of G-Quadruplex-Based Deoxyribozymes

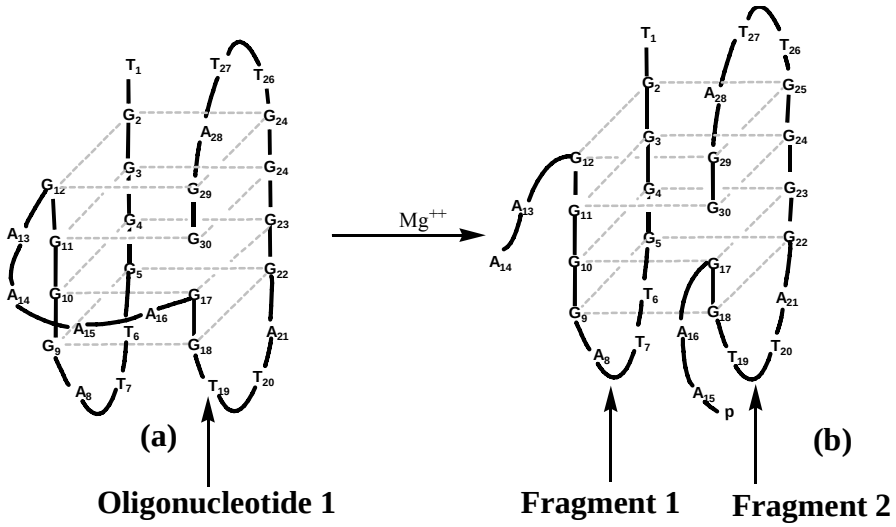
3.1 Background and Aims

Deoxyribozymes, also known as catalytic DNA, are DNA molecules with catalytic action. The first deoxyribozyme was discovered in 1994 by Joyce's group at The Scripps Research Institute.[101] This deoxyribozyme assists in lead ion dependent RNA cleaving operations. Catalytic amplification was found to be 100-fold compared to the uncatalyzed reaction. Many other deoxyribozymes have been developed that catalyze DNA phosphorylation, DNA adenylation, DNA deglycosylation, porphyrin metalation, thymine dimer photoreversion and DNA cleavage. Of particular interest are DNA ligases.[102] These molecules have demonstrated remarkable chemoselectivity in RNA branching reactions. Although each repeating unit in a RNA strand owns a free hydroxyl group, the DNA ligase takes just one of them as a branching starting point. In addition, some deoxyribozymes have found practical use in metal biosensors.[103]

It has been demonstrated that certain catalytic DNA entities such as porphyrin-metalating deoxyribozyme, HD RNA-cleaving DNA enzyme and ATP-utilizing kinase deoxyribozyme utilize G-quadruplex as stable structural cores to sustain their catalytic domains and activities.[104-107] In order to search for new deoxyribozymes on the basis

of G-quadruplex structure, comprehensive studies were carried out during our investigations. It was demonstrated that certain particular G-quadruplex structure would perform site-specific self-cleaving action under certain reaction conditions.[108] In addition, effect of certain factors such as pH value, DNA concentration, metal ions on the self-cleavage processes of G-quadruplex were also carried out. The outcomes of our studies on the new types of deoxyribozymes are discussed in the following sections of this chapter.

3.2 Confirmation of self-cleaving action of a particular G-quadruplex



Oligonucleotide 1: 5' $T_1G_2GGGTTAGGGGAA_{14}A_{15}AGGTTAGGGGTTAG_{29}G_{30}$ 3' (30-mer)
Fragment 1: 5' $T_1G_2GGGTTAGGGGA_{13}A_{14}$ 3' (14-mer)
Fragment 2: 5' p-$A_{15}A_{16}GGTTAGGGGTTAG_{29}G_{30}$ 3' (16-mer)
p: phosphate group

Figure 3-1. Schematic representation of a self-cleavage process of G-quadruplex DNA in this study.

Figure 3-1 depicts a schematic diagram of a DNA self-cleavage processes. A guanine-rich 30-mer oligonucleotide was designed with an expectation that this oligonucleotide would form an externally looped G-quadruplex assembly (a in **Figure**

3-1) under proper conditions. Our initial intention in designing such a guanine-rich oligonucleotide was to examine whether a transesterification reaction could be feasible between the hydroxyl group at its 3' end and the phosphodiester bond between A_{16} and G_{17} since these functional groups are proximal to each other upon G-quadruplex formation. Instead of observing such a designed transesterification reaction, a self-cleavage reaction of **Oligonucleotide 1** at one of the two phosphodiester bonds between A_{14} and A_{15} was observed by coincidence (**Figure 3-1**). This can be suggested that the self-cleaving reaction of **Oligonucleotide 1** took place at the one of the phosphodiester bonds near the 3' end of A14 in the middle of its sequence as illustrated in **Figure 3-2**.

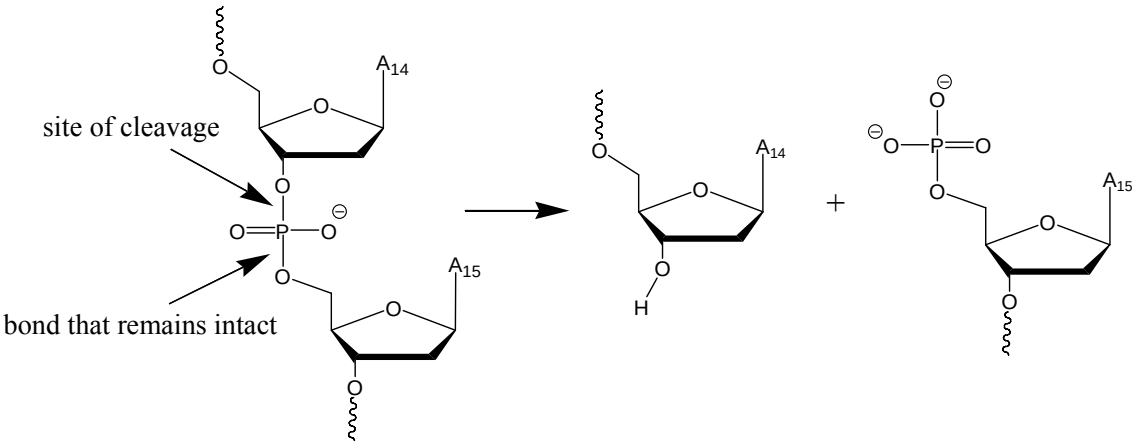


Figure 3-2. Diagrammatic illustration of a possible self-cleaving reaction at one of the two phosphodiester bonds between A_{14} and A_{15} of **Oligonucleotide 1**.

Oligonucleotide 1 was phosphorylated at its 5' end with [γ-^{32}P] ATP in the presence of T4 polynucleotide kinase and further purified by polyarylamide gel electrophoresis and NAP-25 column. In order to allow the formation of proper G-quadruplex assemblies, this guanine-rich oligonucleotide was incubated at room temperature in the presence of 5 mM KCl for 2 hours. The self-cleavage reactions of

Oligonucleotide 1 were initiated by adding Histidine/$MgCl_2$ mixture, which was further kept at 30 °C for a different period of time. As shown in **Figure 3-3**, a new fast moving band was observed when such a reaction was allowed to proceed for 2 hours (band 1 in Lane 2). The mobility shift of this new band is close to that of a molecular weight marker of 14-mer (5' *p-TGGGGTTAGGGGAA 3', Lane 5), which implied that a cleavage reaction took place between A_{14} and A_{15} of this guanine-rich sequence. The yield of the self-cleavage reactions increased with the increase of reaction time and ~50% cleavage of **Oligonucleotide 1** could be achieved within 2 hours.

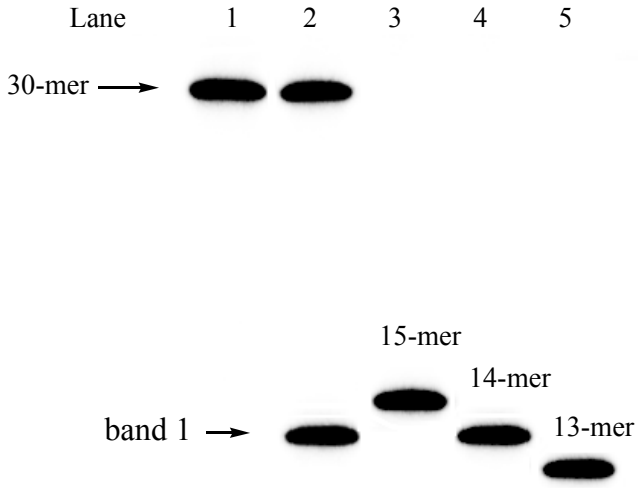


Figure 3-3. Polyacrylamide gel electrophoretic analysis of self-cleavage of DNA. Lane 1: **Oligonucleotide 1** alone; Lanes 2: self-cleavage reactions lasting for 2 h; Lane 3: a 15-mer Oligonucleotide (*p-TGGGGTTAGGGGAAA) alone; Lane 4: a 14-mer (*p-TGGGGTTAGGGGAA) alone; Lane 5: a 13-mer (*p-TGGGGTTAGGGGA) alone.

If a DNA cleavage reaction indeed occurred in the middle of the sequence of **Oligonucleotide 1** in our studies, a second fragment of 16-mer should in theory be generated at the same time. In order to visualize the two fragments simultaneously, SYBER Green staining experiments were conducted next. As shown in **Figure 3-4**, two fast moving bands (Band 1 and Band 2 in Lane 2-5) were visible from the stained polyacrylamide gel, which displayed the same mobility shifts as those of a 14-mer

marker (Lane 6) and a 16-mer marker (Lane 7) respectively. These electrophoretic analysis data are the indications that a cleavage reaction indeed took place between A_{14} and A_{15} in the middle of the sequence of **Oligonucleotide 1** as shown in **Figure 3-1**.

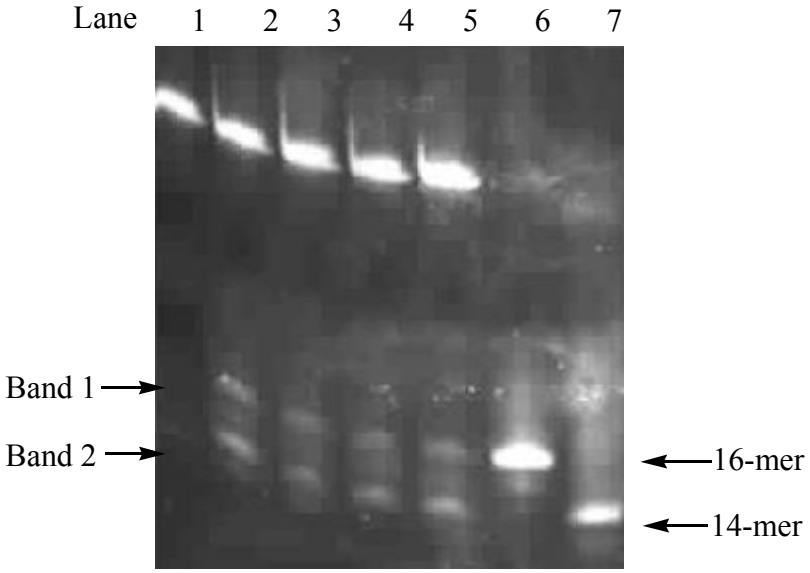


Figure 3-4. Polyacrylamide gel electrophoretic analysis of self-cleavage of DNA visualized by SYBER Green staining. Lane 1: **Oligonucleotide 1** alone; Lane 2-5: self-cleavage reaction lasting for 2 hr (1 μM **Oligonucleotide 1** assay). Lane 6: a 16-mer (5' AAGGTTAGGGGTTAGG 3') alone; Lane 7: a 14-mer (5' TGGGGTTAGGGGAA 3') alone.

3.3 Effect of certain factors on the G-quadruplex-based self-cleavage reaction

3.3.1 Metal ion dependence

One of the distinguishable characteristics of G-quadruplexes from other structural features of nucleic acids is its selective interaction with certain cations that fit in the central cavities formed by guanine tetrads.[109] It has been well established in the past that the preference of monovalent ions by G-quartet in their interactions is in the following order: $K^+ >> Na^+ > Rb^+ > Cs^+ > Li^+$.[109] As was examined, potassium ion is required in this DNA-cleaving reaction. With the intention of knowing how variation

of monovalent ions could affect the self-cleavage reactions of G-quadruplex, K^+ was replaced with Li^+, Na^+, Rb^+ and Cs^+ in this study. As shown in **Figure 3-5**, a self-cleavage product was generated when **Oligonucleotide 1** was incubated with 5 mM K^+ at 20°C for 2 h followed by addition of Mg^{2+}/Histidine to activate the self-cleavage reaction as expected. However, there was no cleavage product observable when K^+ was replaced with Li^+, Na^+, Rb^+ and Cs^+ respectively. The absence of cleavage product in the presence of these four new alkaline ions could be caused by the reason that the G-quadruplex structures containing these four alkali ions might not be able to sustain the right conformation of G-quadruplex needed for the corresponding self-cleavage reaction.

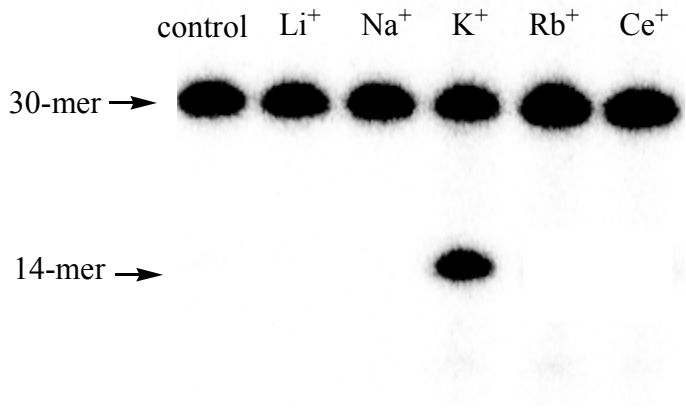


Figure 3-5. PAGE analysis of self-cleavage of **Oligonucleotide 1** in the presence of 5 mM alkaline metal ions (Li^+, Na^+, K^+, Rb^+ and Cs^+)

Besides alkaline metal ions, it is known that certain alkaline earth and transition metal ions could stabilize G-quadruplex assemblies through occupying the central cavity as alkaline ions do.[110] With the aim of examining how these divalent ions affect the self-cleavage reactions of G-quadruplex, K^+ was replaced with Mg^{2+}, Ca^{2+}, Sr^{2+}, Ba^{2+}, Zn^{2+}, Mn^{2+}, Ni^{2+}, Co^{2+} and Pb^{2+}. As shown in **Figure 3-6**, there is no cleavage product generated in the presence of Zn^{2+}, Mn^{2+}, Ni^{2+}, Co^{2+} and Pb^{2+} while a new band was formed between 14-mer and 30-mer when K+ was replaced by Mg^{2+},

Ca^{2+} and Sr^{2+} (**Figure 3-7**). Our subsequently investigation revealed that the fast moving band (band 1 in lane 3, 4 and 5) correspond to an intact G-quadruplex structure rather than a low molecular weight self-cleavage product. This happened most likely because the G-quadruplex structures of **Oligonucleotide 1** containing Sr^{2+} Ca^{2+} and Mg^{2+} were too stable to be disintegrated by denaturing polyacrylamide gel electrophoresis. Interestingly, on the other hand, when K^+ was replace with Ba^{2+}, self-cleavage product of **Oligonucleotide 1** was observed (lane 6 in **Figure 3-7**), which indicates that K+ is not the only metal ion that could sustain the G-quadruplex structure required for next step self-cleavage reactions.

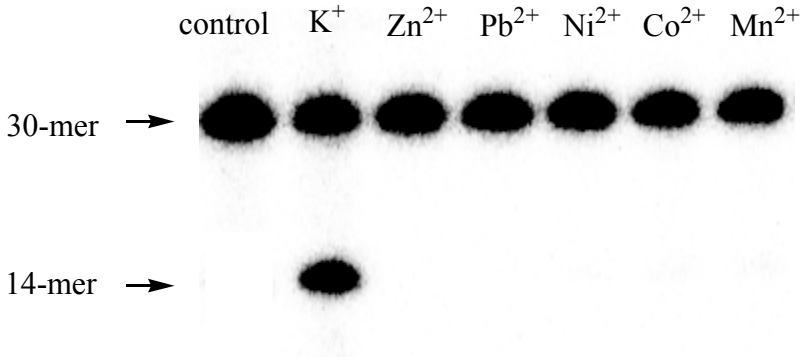


Figure 3-6 PAGE analysis of self-cleavage of **Oligonucleotide 1** in the presence of 1 mM transition metal ions (Zn^{2+}, Pb^{2+}, Ni^{2+}, Co^{2+} and Mn^{2+}).

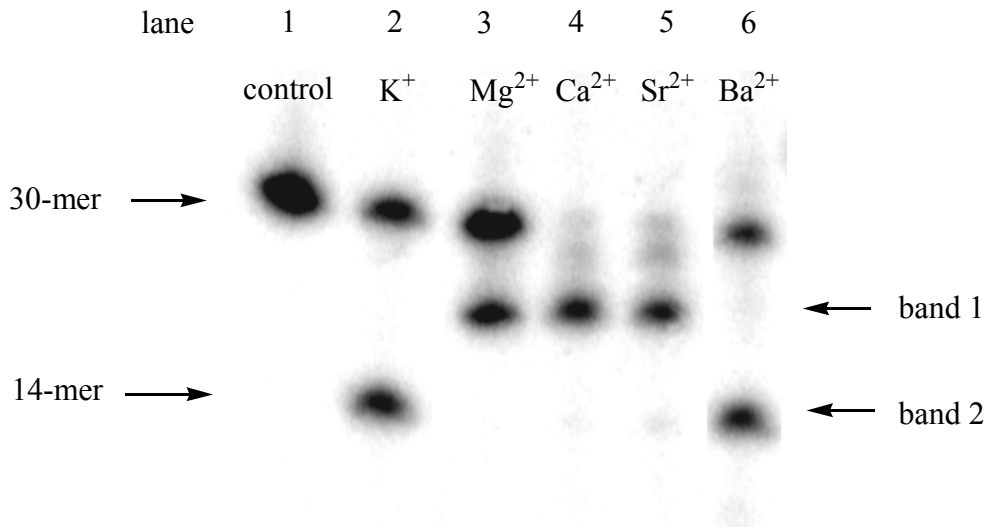
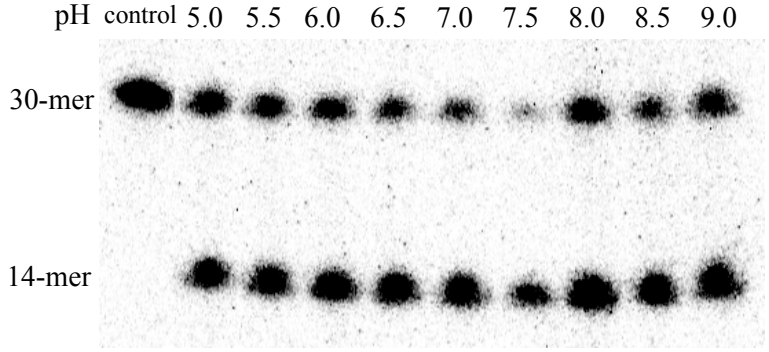


Figure 3-7 PAGE analysis of self-cleavage of **Oligonucleotide 1** in the presence of 5 mM alkaline earth metal ions (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+}).

3.3.2 pH dependence

Catalytic activities of enzymes are commonly affected by the changes of their surrounding pH values.[111] In order to understand the effect of variation of pH value on the reactivity of **Oligonucleotide 1**, the self-cleavage reactions of this guanine-riched oligonucleotide at certain different pH were accordingly carried out. As seen in **Figure 3-8**, the maximum yield of the self-cleavage reactions occurred in the pH range from 6.5 to 7.5. As shown in **Figure 3-8b**, when pH value was set to above 7.5 or below 6.5, a phenomenon similar to many enzymatic processes. On the other hand, the cleavage reaction rate of **Oligonucleotide 1** decreased when pH value is above 7.5 and below 6.5. Since the variation of pH values has little effect on the stability of G-quadruplex,[111] the low efficiency of the self-cleavage reaction of **Oligonucleotide 1** could be due to certain subtle conformational change of G-quadruplex at pH beyond the range of 7.5 to 6.5 that disfavor the catalytic process.

a)



b)

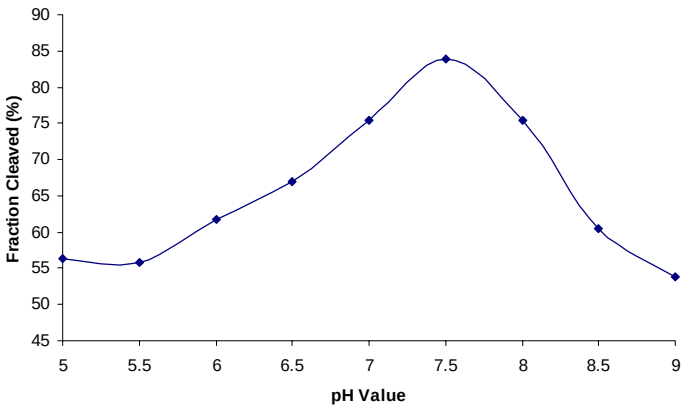


Figure 3-8 pH dependent of self-cleavage of **Oligonucleotide 1** vary from 5.0 to 9.0. The same procedures as those for preparing samples loaded in lane 2 in **Figure 3-3** were used except that pH value of buffer in the new experiments varied.

3.3.3 DNA concentration dependence

Earlier investigations demonstrated that the conformation of G-quadruplex was affected by the concentration of DNA. With the aim of finding out whether the concentration of **Oligonucleotide 1** affects our designed self-cleavage course, we examined **Oligonucleotide 1** vary from 1 nM to 1000 nM. As shown in **Figure 3-9**, there was trace of DNA cleavage detectable when concentration of **Oligonucleotide 1** is higher than 100 nM when cleavage reactions lasted for 2 h. This most likely due to a different template of G-quadruplex generated before the cleavage course under high concentration of **Oligonucleotide 1** or longer reaction time needed.

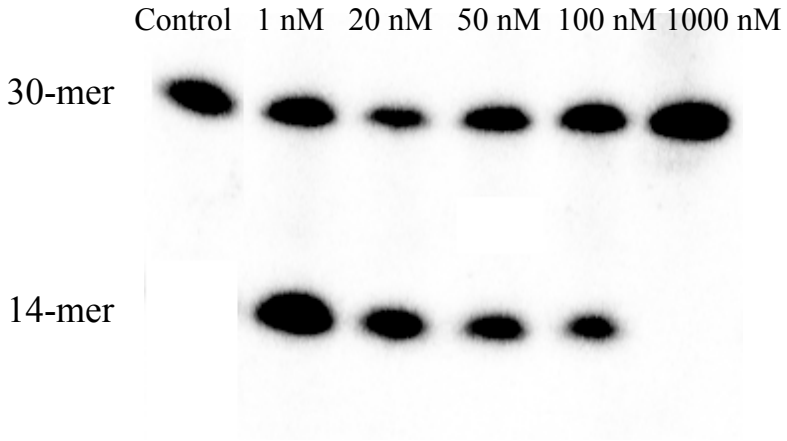


Figure 3-9. PAGE analyses of self-cleavage of **Oligonucleotide 1** in different DNA concentrations vary from 1 nM to 1000 nM. The same procedures as those for preparing samples loaded in lane 2 in **Figure 3-3** were used except that concentration of **Oligonucleotide 1** in the new experiments varied.

3.3.4 Determination of rate constants of the G-quadruplex-based self-cleavage reactions

Time dependence of these self-cleavage reactions was also carried out in our studies. As shown in **Figure 3-10**, the yield of the self-cleavage reactions increased

with the increase of reaction time and ~50% cleavage of **Oligonucleotide 1** could be achieved within ~2 hours when *Mg^{2+}/histidine* complex was used to trigger the reaction.

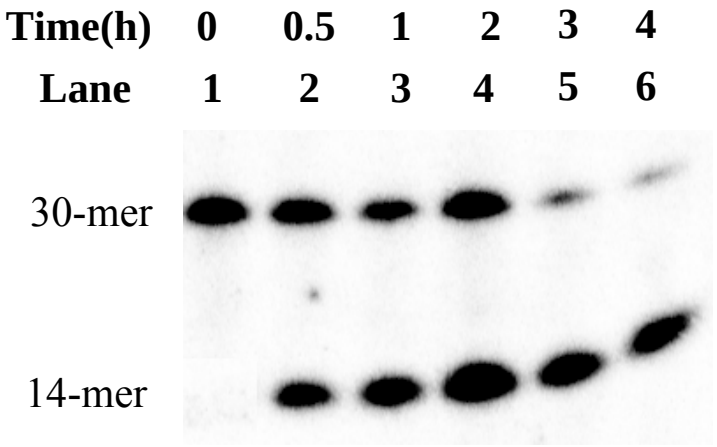


Figure 3-10. Time dependence of self-cleavage reaction of **Oligonucleotide 1**. The same procedures as those for preparing samples loaded in Lane 2 in **Figure 3-3** were used except that the reactions were stopped at different time intervals. The time of reaction in Lanes 1, 2, 3, 4, 5 and 6 were 0, 0.5, 1, 2, 3 and 4 hours respectively.

Similar to the actions of many cis-acting ribozymes and deoxyribozymes, self-cleavage reaction of **Oligonucleotide 1** appears to be a unimolecular process. Rate constants of the self-cleavage reactions of **Oligonucleotide 1** were accordingly determined on the basis of first order kinetics[112] during our investigations. The obtained maximum rate constant under our standard condition is 0.032 min^{-1} (30 °C pH 7.2) (**Figure 3-11**) which is close to those of certain hairpin ribozymes (0.1 min^{-1} to 0.3 min^{-1}) [113] in their self-cleaving courses of action. Given that the rate constant of spontaneous hydrolysis of DNA in the absence of catalyst is ~$10^{-12}min^{-1}$,[114] **Oligonucleotide 1** produces in effects ~ 3.2 x 10^{110} (0.032 $min^{-1}/10^{-12}$ min^{-1}) fold rate enhancement in its DNA cleavage reaction. The rate enhancement generated by hairpin ribozymes is only 3 x 10^{8} (0.3 $min^{-1}/10^{-7}$ min^{-1}) since a rate constant of spontaneous hydrolysis of RNA in the absence of catalyst (~10^{-7} min^{-1}) [115] is much faster than that of DNA. In view of the extremely inert nature of phosphodiester

bonds of DNA as well as high degree of rate enhancement attained by **Oligonucleotide 1**, it is our opinion that this newly discovered hydrolytic self-cleaving deoxyribozyme could be taken as an efficient biocatalyst.

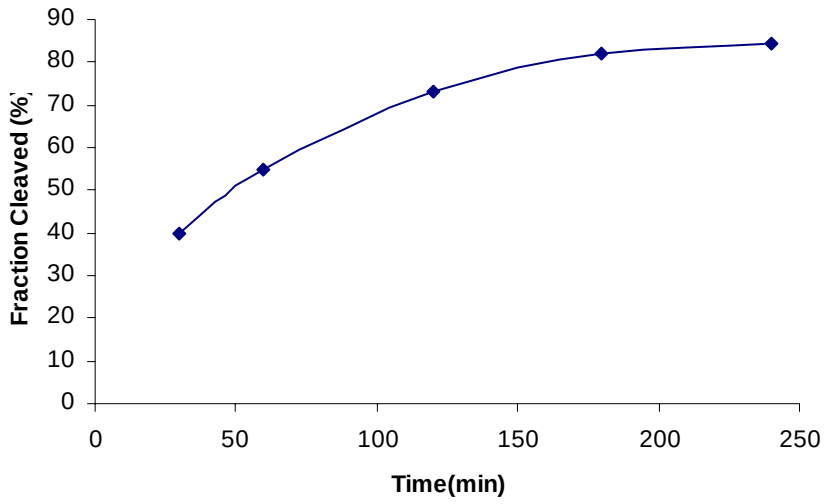


Figure 3-11. Determination of observed rate constants of **Oligonucleotide 1** in its self-cleavage reactions. Plot of fraction of Fragment 1 versus reaction time. The observed rate constant (k_{obs} = 0.032 min^{-1} at 30 °C) was obtained by fitting the corresponding data to single exponential equation $F = F_{max} \times (1+ e^{-kt})$, where F is the fraction of Fragment 1 at time t.

On the other hand, in order to determine whether the presence of magnesium ions is indeed a prerequisite for the self-cleavage process, a reaction mixture (5mM pH 7.0 HEPES/Histidine, 20 mM KCl and **Oligonucleotide 1**) containing no magnesium ion was examined during our investigations. As shown in **Figure 3-12**, self-cleavage reactions of **Oligonucleotide 1** was only observable after the corresponding reaction mixtures were incubated > 42 h and yield of the self-cleavage reaction could reach 50.6% in 56 h (92.6% in 72 h). These observations are the indication that self-cleavage reaction of **Oligonucleotide 1** is still capable of taking place in the absence of magnesium ions.

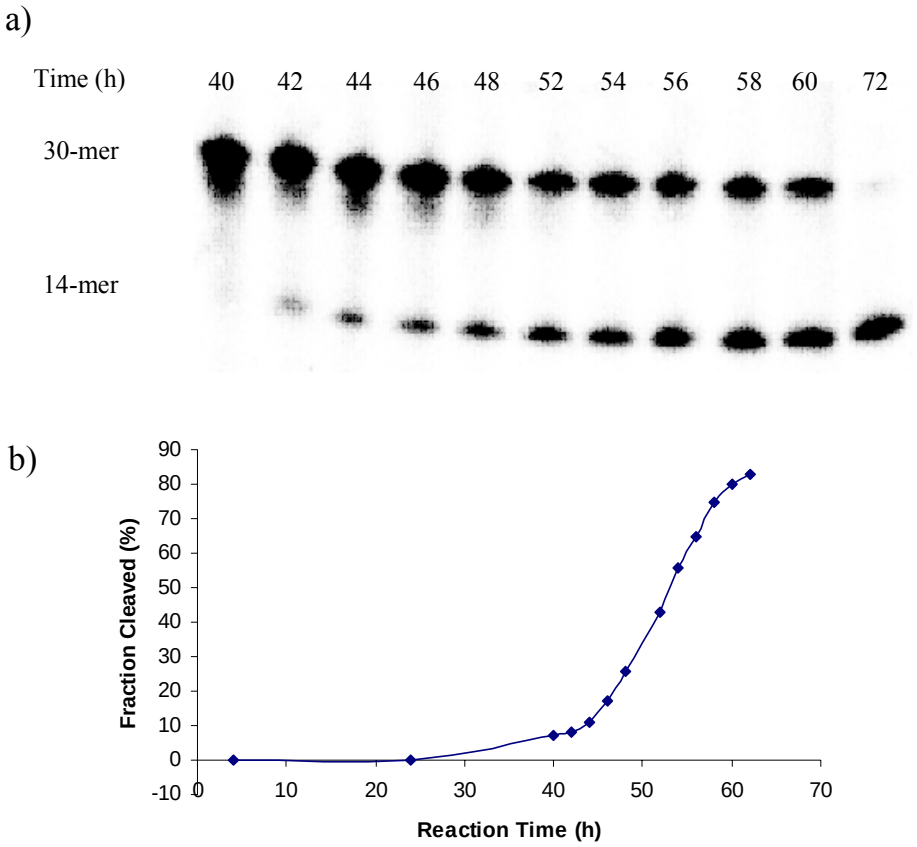


Figure 3-12 Time dependence of self-cleavage reaction of **Oligonucleotide 1** in the absence of magnesium ions. The sample was kept in a mixture of 5mM pH 7.0 HEPES/Histidine, 20 mM KCl and stopped at different time intervals varied from 24h to 58 h.

3.3.5 Potassium ion concentration dependence

As we discussed in Chapter 3.3.1, potassium ion is one of the preferable monovalent cations for stabilizing G-quadruplex structures of DNA.[109, 116-118] Potassium concentration dependence of the self-cleavage reaction was further examined. As shown in **Figure 3-13**, the yield of DNA cleavage product decreased when the concentration of potassium ion increased as high as 100 mM. This happened most likely because that high concentration of potassium ion probably could induce another conformation of G-quadruplex structure, which disrupts the self-cleavage reactions.

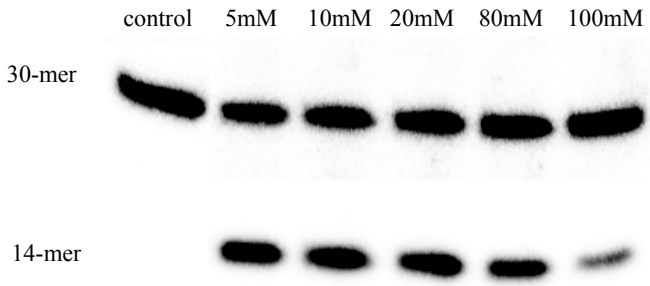


Figure 3-13. Effect of potassium ion concentration on the self-cleavage reaction of **Oligonucleotide 1**. The same procedures as those for preparing samples loaded in lane 2 in **Figure 3-3** were used except that concentration of potassium chloride in the new experiments varied.

It was discovered earlier that high concentration of sodium ion could sustain a different conformation of G-quadruplex structure from those in the presence of low concentration of Na^+. Self-cleavage reactions of **Oligonucleotide 1** in the presence of high concentration of monovalent ions were accordingly carried out. As shown in **Figure 3-14**, **Oligonucleotide 1** was incubated with 80 mM of Li^+, Na^+, K^+, Rb^+ and Ce^+ ions separately for 2 h before the self-cleavage reaction was triggered. Surprisingly, besides potassium, sodium and rubidium ion also generate a small amount of DNA cleavage product when *Mg^{2+}/Histidine* complex was added. This happended most probably because the anti-parallel G-quadruplex structure (**Figure 3-1**) needed for our self-cleavage reaction could also be sustained by high concentration of sodium or rubidium ion.

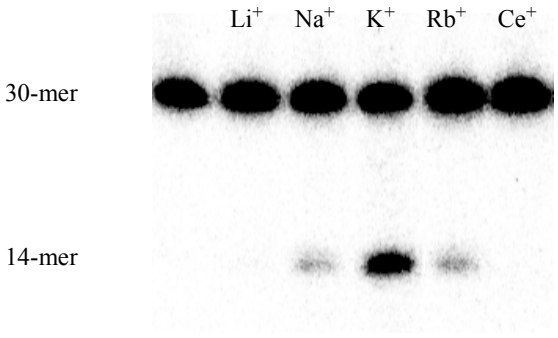


Figure 3-14 PAGE analysis of self-cleavage of **Oligonucleotide 1** in the presence of 80 mM alkaline metal ions (Li^+, Na^+, K^+, Rb^+ and Cs^+).

3.3.6 The formation of G-quadruplex by Oligonucleotide 1

To verify the formation of G-quadruplex structure by **Oligonucleotide 1**, CD spectroscopic analysis on a solution containing this guanine-rich sequence was also carried out. As shown in **Figure 3-15**, this **Oligonucleotide 1**-containing solution displayed a maximum absorption at 295 nm at 34 °C, which is a characteristic sign of the presence of anti-paralleled G-quadruplex structures in the mixture.[119-121] When the **Oligonucleotide 1**-containing solution was heated up to 90 °C, the disappearance of a positive peak at 295 nm represented that the anti-parallel structure is no longer exist.

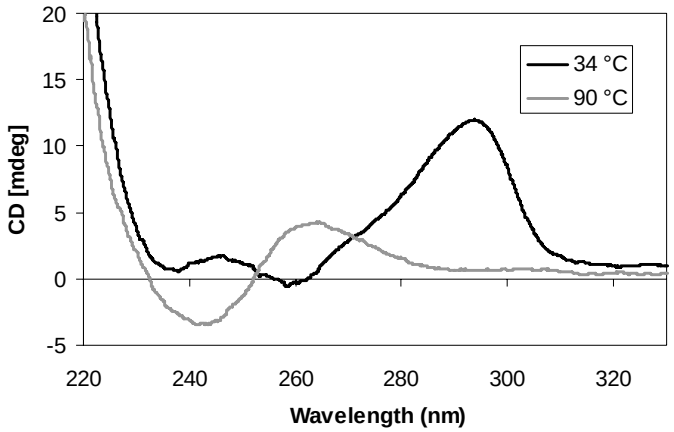


Figure 3-15. CD spectroscopic analysis of **Oligonucleotide 1** in the presence of K^+. A mixture (pH 7.0) containing 5 mM HEPS, 5 mM NaCl, 5 mM KCl and 10 μM **Oligonucleotide 1** was examined with a CD Spectropolarimeter at 34 ^oC (black) and 90 ^oC (grey) respectively over an range of wavelengths from 220 nm to 330 nm.

CD spectroscopic examination on **Oligonucleotide 1**-containing solutions in the presence of the other four monovalent ions (Li^+, Na^+, Rb^+ and Ce^+) were conducted during our investigations. As it can be seen in **Figure 3-16**, strong positive peaks near 260 nm were observed when **Oligonucleotide 1** was incubated with Li^+, Na^+, Rb^+ and Ce^+ which is the sign of presence of parallel conformation instead of anti-parallel type.

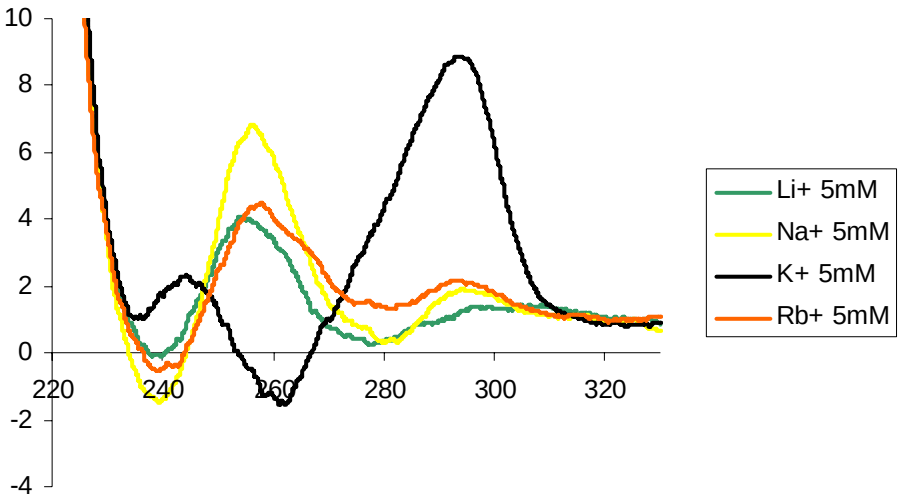


Figure 3-16. Comparison CD studies of **Oligonucleotide 1** in the presence of different alkaline metal ions (Li^+, Na^+, K^+ and Rb^+).

Similar phenomenon were observed when K^+ was replaced by divalent ions (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+} tested), while barium ion can form major ant-parallel structure and consequently showed a positive peak at 295 nm (**Figure 3-17**). It is interesting that the spectroscopic data are consistent with the results generated from the studies of ion-dependent of self-cleavage reaction as shown in **Figure 3-7**.

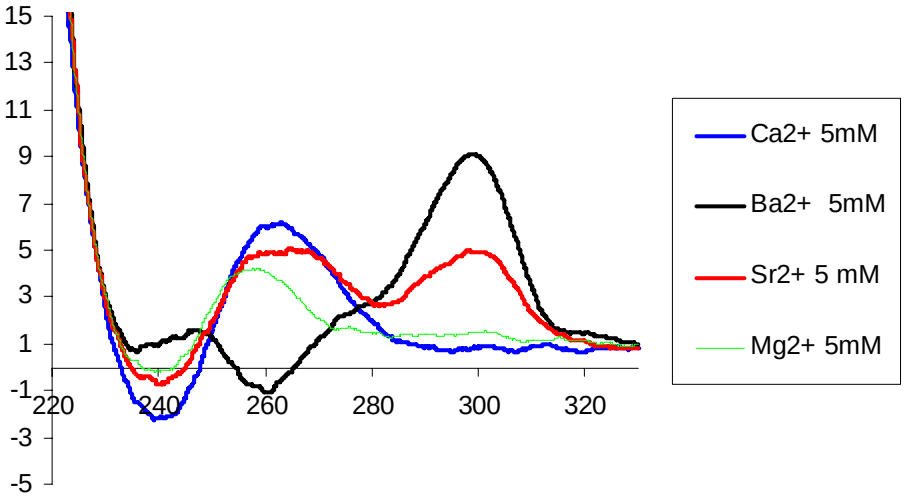


Figure 3-17. Comparison CD studies of **Oligonucleotide 1** in the presence of different alkaline earth metal ions (Mg^{2+}, Ca^{2+}, Sr^{2+} and Ba^{2+}).

On the other hand, the CD spectrum of **Oligonucleotide 1**-containing solution in the presence of strontium ion exhibited dual peaks with equal intensity at 260 nm and 296 nm respectively, which indicates that this ion could sustain both parallel and anti-parallel conformation of G-quadruplex simultaneously, which could be the cause of absence of self-cleavage product when this divalent ion was used. In addition, CD spectrums of **Oligonucleotide 1** containing certain transition metal ions such as Zn^{2+}, Pb^{2+}, Ni^{2+}, and Mn^{2+} do not give any peak with reasonable high intensity at 295 nm (**Figure 3-18**), which is also consistent with the results of our self-cleavage studies shown in **Figure 3-6**.

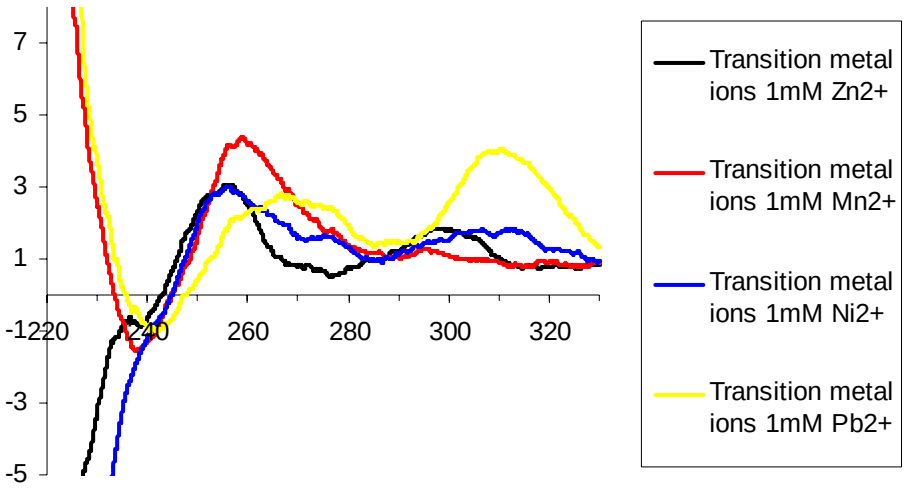


Figure 3-18. Comparison CD studies of **Oligonucleotide 1** in the presence of different transition metal ions (Zn^{2+}, Pb^{2+}, Ni^{2+}, and Mn^{2+}).

3.4 Conclusions

Self-cleavage reaction of a guanine-riched oligonucleotide (**Oligonucleotide 1**) was thoroughly studied during our investigations. It appeared that presence of potassium ions is a prerequisite of the self-cleavage reaction while barium ion can also sustain the particular G-quadruplex structure for further self-cleavage reaction. The results of our studies are consistent with the suggestion that G-quadruplex is the core structure needed for the self-cleavage process. More interestingly, our kinetic studies revealed that the self-cleavage rate of **Oligonucleotide 1** could be as fast as certain hairpin ribozymes.[113] In addition, the site-specific hydrolysis reaction could still take place when in the absence of magnesium ion as long as the corresponding reaction mixture was kept for 2 days. It is our hope that the results reported in the current studies could be helpful for searching for new G-quadruplex structures that could perform self-cleavage reactions.

Chapter 4

Construction of Fluorescein-Tagged Circular G-Quadruplexes

4.1 Background and Aims

The structural feature of G-quadruplex is a self-assembling entity of DNA composed of two or more stacks of G-quartets and formed either via either unimolecular assembling of DNA or intermolecular association.[122,123] The adjustable nature of strand stoichiometry[124-137] allows this self-assembling entity of G-quadruplex to be formed by association of one (unimolecular), two (bimolecular) or four (tetramolecular) oligomeric molecules. The well-investigated examples of the linear oligonucleotides that constitute unimolecular, bimolecular and tetramolecular structures of G-quadruplex include d($G_3T_2AG_3T_2AG_3T_2AG_3$) (3.5 human telomere), [d($G_4T_4G_4$)]$_2$, (Oxy-1.5) and [d(TGGGGT)]$_4$ respectively.[132, 138-144]

It was demonstrated previously in our group[145] that certain bimolecular structures of G-quadruplex were capable of acting as templates for directing the formation of circular oligonucleotides with high efficiency and sequence specificity. One of the intriguing aspects of these bimolecular G-quadruplex entities as templates was its mutual templating fashion on the scale of individual oligonucleotides in which one linear strand serves as a template for another. This mutual-templating pathway

consequently led to a unique G-quadruplex entity constituted of two identical circular oligonucleotides.[145]

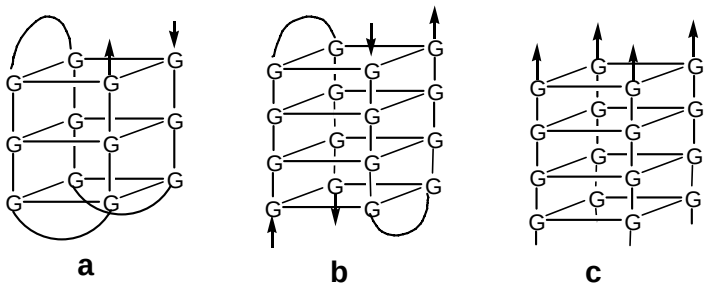


Figure 4-1. Schematic representation of G-quadruplex formed unimolecularly (***a***), bimolecularly (***b***) and through the association of four strands of oligonucleotides (***c***).

On the other hand, a variety of proteins have been discovered that are capable of binding to certain structural forms of G-quadruplexes. Some aptamers such as an inhibitor of HIV integrase[146] and the thrombin-binding aptamer [147] are formed through the utilization G-quadruplex as their structural cores. Moreover, several G-quadruplex-binding proteins have been identified over the past ten years such as rat liver proteins uqTBP25.[148] On the basis of earlier discoveries,[146-151] construction of unimolecularly circular G-quadruplexes tagged with fluorescent molecules as molecular probes was accordingly set up as one of our long-term research objectives. One of the advantages in our selection of circular oligonucleotides as diagnostic probes is their complete resistance to exonuclease hydrolysis, a property resulted from the absence of open terminus within their structures.[152] In addition, the structured DNA composed of circular sequence often possesses a higher melting point than those constituted of linear components while higher stability is always preferred in the design of molecular probes.[153] Therefore, we anticipate that the fluorescence-labeled circularly unimolecualr G-quadruplexes could act as useful probes for identifying new types of enzymes and other functional proteins correlated with G-quadruplex *in vivo*

and for elucidating the binding patterns between G-quadruplex and its corresponding binding proteins. The outcomes of our investigation on design, synthesis and characterization of the particular circular oligonucleotides on the template basis of G-quadruplex structures are discussed in the following sections of this chapter.

4.2 Construction of Circular Oligonucleotides on the Basis of Unimolecular G-Quadruplex

4.2.1 Design and Synthesis of Circular Oligonucleotides on the Basis of Unimolecular G-Quadruplex

In the initial step, exploration of the feasibility of constructing unimolecular G-quadruplex without a fluorescence tag (f in **Figure 4-2**) was carried out. As expected, it is demonstrated that certain unimolecularly circular G-quadruplexes can indeed be attained from linear precursors through chemical ligation reactions. In addition, in consideration of the fact that several new proteins discovered during the past few years facilitate either the formation of linear unimolecular G-quadruplex or disintegration of it,[154-160] it is our anticipation that the unimolecularly circular G-quadruplexes attained in our studies (f in **Figure 4-2**) could be a unique tool for unveiling the winding and resolving mechanisms of G-quadruplex mediated by these G-quadruplex-promoting and disintegrating proteins.

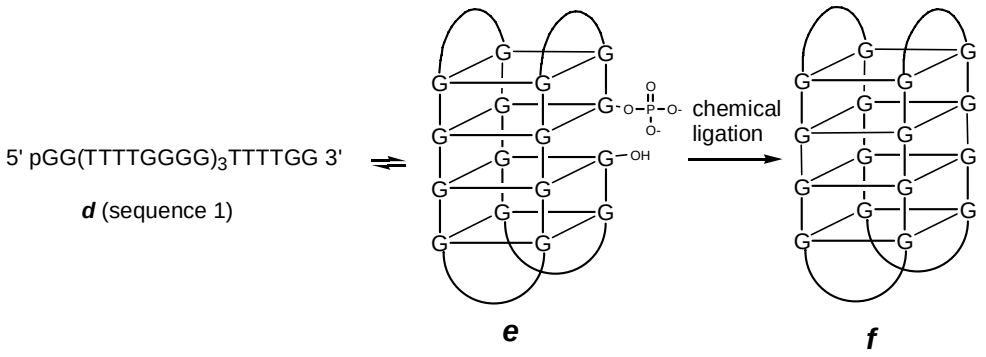
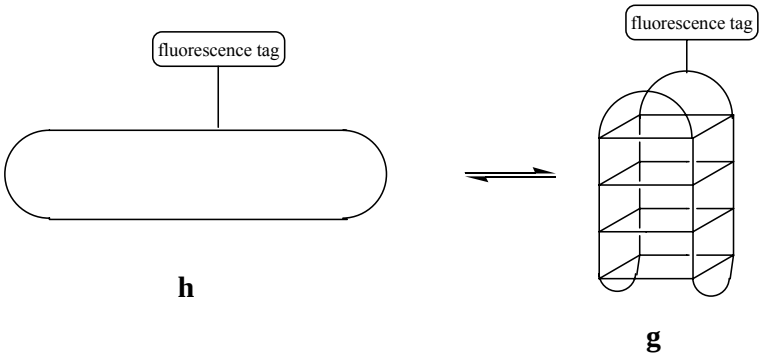


Figure 4-2. Diagrammatic illustration of our strategy for constructing unimolecularly circular G-quadruplex through chemical ligation.

A unimolecular G-quadruplex template was selected (g in **Figure 4-3**) rather than its bimolecular counterpart (i) as the synthetic target in our strategy. It is because that the concentrations of diagnostic molecules in many practical applications are very low[161, 162] while bimolecular complexes (i), once dissociated into the corresponding unimolecular components (j), might not readily go back to its original state at a very low concentration due to kinetic unfavorableness. In addition, a fluorescence-tagged bimolecular G-quadruplex can in theory possess two isomeric forms (i and k) while one of them (k) may not be suitable as a molecular probe because of the presence of fluorescence tags at both ends of its columnar structure that prevent the approach of G-quadruplex-interacting proteins. This co-existence of two isomeric forms could decrease the efficiency of bimolecular G-quadruplexes as molecular probes in practical application. In contrast, unimolecular G-quadruplex could, even if it is denatured occasionally, readily return to its original folding state as the result of kinetic favorableness. It consequently appears that selection of unimolecular G-quadruplexes as molecular probes has certain advantages over the choice of their bimolecular counterparts.



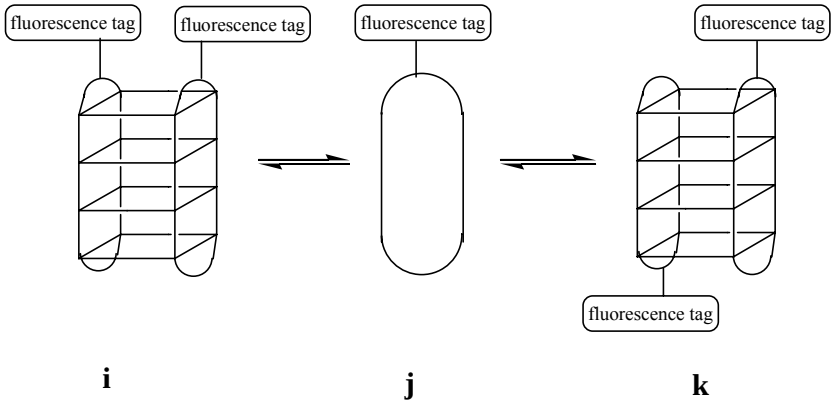


Figure 4-3. Possible folding patterns of certain fluorescence-tagged circular G-quadruplexes.

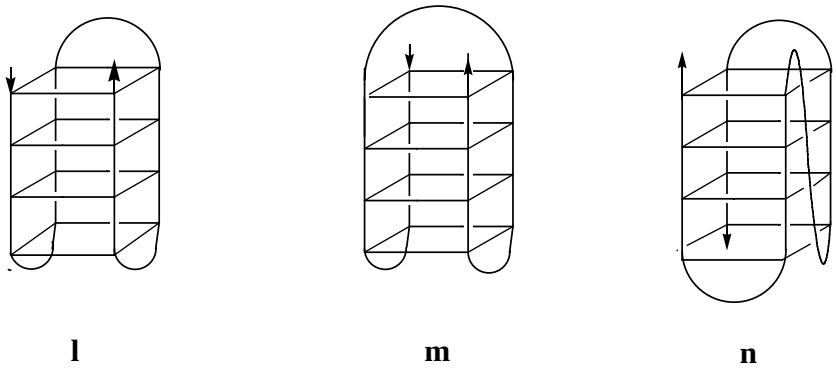


Figure 4-4. Illustration of different loop geometries possessed by unimolecular G-quadruplexes.

The loops in a unimolecular G-quadruplex can in general be connected in three different ways (**Figure 4-4**). It appears that the only conformation that can be utilized to synthesize circular oligonucleotides is the one with its loops connected in adjacent-adjacent-adjacent fashion (l in **Figure 4-4**). **Figure 4-2** accordingly illustrates our strategy for manipulating the unimolecular entity of G-quadruplex to facilitate the construction of guanine-rich circular oligonucleotides. A linear sequence, 5' $GG(TTTTGGGG)_3TTTTGG3'$ (**Sequence 1**), was selected as the precursor of our circularization reaction. It is anticipated that under certain conditions this linear oligonucleotide would self-assemble into a desired unimolecular complex,

within which its two termini are forced to align in a proximal position guided through reversed Watson-Crick interaction.[163] Once this self-assembling entity (e in **Figure 4-2**) is generated, subsequent chemical ligation reaction between the two open termini would lead to a desired circular G-quadruplex constituted of a single strand of oligonucleotide.

Sequence 1 was accrodingly incubated in a pH 6.0 buffer to allow the desired structure of G-quadruplex to generate (d to e in **Figure 4-2**). The 5' terminal phosphate within the unimolecular complex was following activated by the addition of N-cyanoimidazole to facilitate the formation of a phosphodiester bond with its adjacent 3' hydroxyl group (e to f in **Figure 4-2**). As shown in **Figure 4-5**, a new product (Lane 9) was generated from this ligation reaction with its rate of mobility shift slower than that of **Sequence 1** (Lane 1). It was demonstrated earlier by Wang and Kool[164] that a 34-mer circular oligonucleotide moved slower than that of its corresponding linear precursor. Based on this information, the newly formed product from our ligation reaction was identified tentatively as the circular oligonucleotide generated from its corresponding linear precursor, **Sequence 1**. The yields of this ligation reactions was found to be dependent on reaction time: circularization of the linear precursor proceeded in 13, 29, 42, 59, 67 and 71% yield when the corresponding reactions continued for 5 (Lane 4), 30 (Lane 5), 60 (Lane 6), 120 (Lane 7), 240 (Lane 8) and 360 min (Lane 9) respectively and the overall yield of the ligation reaction is 39% after gel purification.

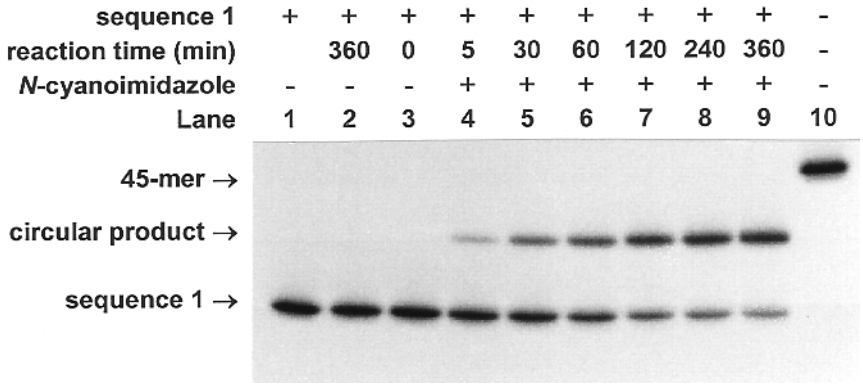


Figure 4-5. Construction of unimolecularly circular oligonucleotides on the template basis of G-quadruplex and time course of the ligation reaction (refer to Chapter 6 for Chemical ligation reactions protocols). The reaction times for these ligation reactions were 0 (Lane 3), 5 (Lane 4), 30 (Lane 5), 60 (Lane 6), 120 (Lane 7), 240 (Lane 8) and 360 min (Lane 9) respectively. Lane 1: **Sequence 1** alone. Lane 2: same reaction as the one loaded in Lane 9 except for the absence of N-cyanoimidazole. Lane 10: the 45-mer of 5'$T_3CGT_2CGTGCT_2GCT_3GCTCT_2G_3T_2AG_5T_2AG_2T_4$3' as a molecular weight marker.

It should be noticed that the circular oligonucleotide possessing three guanines in a row in its sequence，<TTAGGGTTAGGGTTAGGGTTAGGG>，can not be synthesized using the same methodology as discussed above starting from a linear precursor of 5' GGTTAGGGTTAGGGTTAGGGTTAG 3' or 5' GTTAGGGTTAGGGTTAGGGTTAGG 3' while these two linear oligonucleotides are closely correlated with the sequences of telomeres of higher organisms. In addition, the type of circular oligonucleotides synthesized in our studies can form only the structure of G-quadruplex with its loop connected in an adjacent-adjacent-adjacent fashion (l in **Figure 4-4**). The G-quadruplex of biological importance in many cases are, however, the ones exhibiting parallel orientation (c in **Figure 4-1**) or possessing loops connected in adjacent-diagonal-adjacent manner (m in **Figure 4-4**)[165-171] while the structural feature possessing adjacent-adjacent-adjacent loops (l) are presumably present only in a limited number of biological processes.[172, 173] These disadvantages could accordingly prevent the use of unimolecularly circular G-quadruplexes in a

wider range of practical application.

4.2.2 Confirmation of Circular Nature of Our Ligation Product

Exonuclease VII is an exonuclease that hydrolyzes nucleotides from both 3' and 5' ends of single-stranded linear sequence of deoxyribonucleic acids. Circular oligonucleotides are known to resist the degradation by this enzyme[153, 174] due to the absence of open termini within their structures. In order to confirm the circular nature of its phosphate-sugar backbones, the newly formed product from our ligation reaction was purified via denaturing polyacrylamide gel electrophoresis and then hydrolyzed by this exonuclease. As shown in **Figure 4-6**, there was absent of degradation product observable from the reaction of our identified circular product with exonuclease VII (Lane 4). As a positive control, the hydrolysis reaction of **Sequence 1** with exonuclease VII was also carried out, which consequently led to lower molecular-weight products in quantitative yield (Lane 2). The complete resistance of our ligation product to the hydrolysis by this exonuclease should be considered as the indication that there is absent of open terminus within its macromolecular structure.

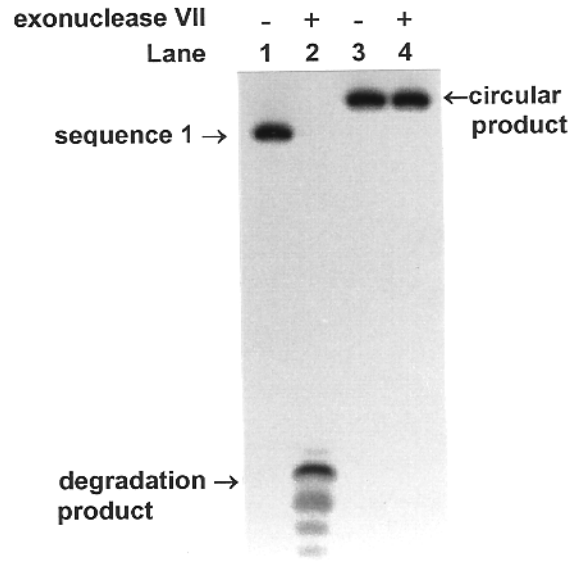


Figure 4-6. Hydrolysis of the identified circular products by exonuclease. Lane 1: **Sequence 1** alone. Lane 2: **Sequence 1** after treatment with 20 units of exonuclease VII at 37 °C for 2 h. Lane 3: the identified circular products alone. Lane 4: the identified circular product after treatment with 20 units of exonuclease VII at 37 °C for 2 h.

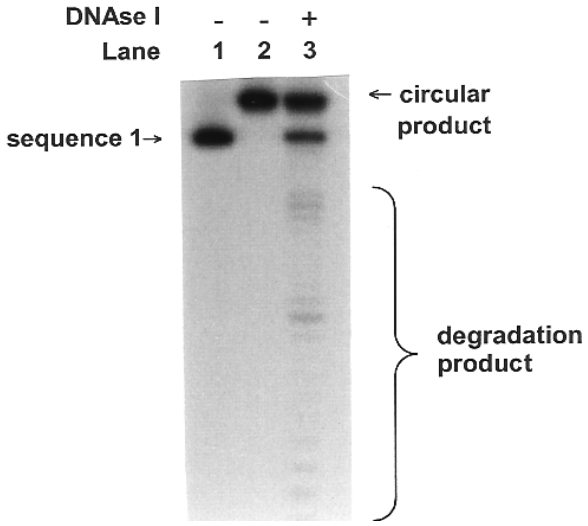


Figure 4-7. Partial hydrolysis of the identified circular products by DNAse I. Lane 1: **Sequence 1** alone. Lane 2: the identified circular products alone. Lane 3: the identified circular product after treatment with 10 units of DNAse I at 25 °C for 30 min.

Previous gel electrophoretic studies by others[164,174] revealed that partially randomized hydrolysis of circular sequences of oligonucleotides resulted in a new major band with its rate of mobility shift nearly equal to that of its linear precursor. The reason leading to this observation is that cleavage of any of the phosphodiester bonds within a circular structure will give rise to the corresponding linear products possessing identical molecular weight.[174] With the aim of further confirming the circular nature of our ligation product, partial hydrolysis tests were carried out during our investigation through the use of DNAse I, an enzyme that hydrolyze both single and duplex forms of DNA in a randomized pattern. As shown in **Figure 4-7**, a major band was observed upon partial hydrolysis of our ligation product that exhibited the same rate of mobility shift as that of its linear precursor (Lane 3). In addition, there was absent of intermediate bands between our ligation product (higher band in Lane 3) and the major band produced from it, which was the characteristic pattern of randomized partial hydrolysis of circular oligonucleotides by enzyme or chemicals.[164, 174]

4.2.3 Conformation Dependence of the Circularization Reactions

If the designed self-assembly of G-quadruplex (e in **Figure 4-2**) is indeed an intermediate structure of our ligation reactions, any alteration from this conformation should have an effect on our circularization courses. In order to examine the effect of this conformation dependency, four new sequences were designed during our investigations in which one, two, three or four non-guanine nucleotides appeared at their 5' or 3' end. As shown in **Figure 4-8**, none of the four mismatched sequences exhibited any observable amount of circularization product (Lanes 3 to 10) under the same reaction condition as the one designed for **Sequence 1** (Lane 2). These

experimental results indicated that a single-base variation led to a distorted conformation of G-quadruplex that was consequently incapable of sustaining correct proximity between the 5' and 3' termini of the same sequence as required for further chemical ligation reaction.

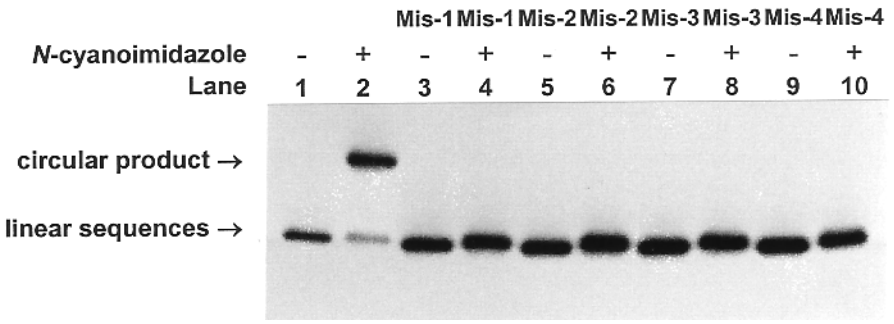


Figure 4-8. Effect of mismatched sequences on the circularization reaction. Ligation reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for replacing **Sequence 1** with the corresponding mismatched sequences. Lane 1: **Sequence 1** alone; Lane 2: reaction mixture of **Sequence 1**; Lane 3: mis-1 (5' **T**$GT_4G_4T_4G_4T_4G_4T_4G_2$ 3') alone; Lane 4: reaction mixture of mis-1; Lane 5: mis-2 (5' **T**$GT_4G_4T_4G_4T_4G_4T_4G$**T** 3') alone; Lane 6: reaction mixture of mis-2; Lane 7: mis-3 (5' **TT**$T_4G_4T_4G_4T_4G_4T_4G$**T** 3') alone; Lane 8: reaction mixture of mis-3; Lane 9: mis-4 (5' **TT**$T_4G_4T_4G_4T_4G_4T_4$**TT** 3') alone; Lane 10: reaction mixture of mis-4.

In order to further probe the effect of sequence variation on our circularization reaction, four additional linear oligonucleotides were selected that were shorter than **Sequence 1** by one (rec-1), two (rec-2), three (rec-3) or four (rec-4 in **Figure 4-9**) nucleotide units. Our expectation from the design of these new sequences was that G-quadruplex, once it was formed from rec-1, rec-2, rec-3 or rec-4, would possess recessive ends in the middle of its columnar structure. As shown in **Figure 4-7**, a small amount of circularization product was generated when the ligation reaction began with rec-1 (Lane 4) while there was absent of circular product observable in the reactions where the sequences of rec-2 (Lane 6 in **Figure 4-9**), rec-3 (Lane 8) and rec-4 (Lane 10) were used as reactants. These observations could be the indication that a

proper structural feature of G-quadruplex can still be generated from a linear oligonucleotide shorter than **Sequence 1** by one nucleotide unit, but not by the others that are two or more nucleotide units shorter.

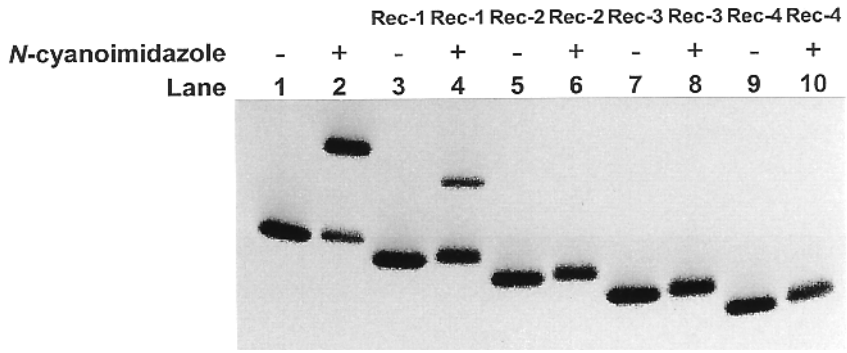


Figure 4-9. Effect of recessive sequences on the circularization reaction. Reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for replacing **Sequence 1** with the corresponding recessive sequences. Lane 1: **Sequence 1** alone; Lane 2: reaction mixture of **Sequence 1**. Lane 3: rec-1 (5' $GT_4G_4T_4G_4T_4G_4T_4G_2$ 3', 31-mer) alone. Lane 4: reaction mixture of rec-1; Lane 5: rec-2 (5' $T_4G_4T_4G_4T_4G_4T_4G_2$ 3', 30-mer) alone; Lane 6: reaction mixture of rec-2; Lane 7: rec-3 (5' $T_4G_4T_4G_4T_4G_4T_4G$ 3', 29-mer) alone ; Lane 8: reaction mixture of rec-3; Lane 9: rec-4 (5' $T_4G_4T_4G_4T_4G_4T_4$ 3', 28-mer) alone ; Lane 10: reaction mixture of rec-4.

4.2.4 Loop-Size Dependence of Our Circularization Reactions

One of the distinctive features of G-quadruplex in structure is its possession of connecting loops that link the guanine tracts within the unimolecular and bimolecular complexes of G-quadruplexes.[175] The sequences of $d(G_4T_nG_4)$ ($n \geq 2$), for example, are capable of forming looped quartet structures whose stabilities increase with the increase in the number of intervening thymine residues within their loops.[132, 177-179] When a single thymine nucleotide is present between the guanine tracts within a linear oligonucleotide such as $d(G_4TG_4)$, however, this guanine-rich sequence can only adopt parallel four stranded structure rather than an antiparalell looped dimeric complex.[177] This took place because the backbone of a single thymine residue is too

short to span over the two edges of a G-quadruplex structure. In consideration of the fact that connecting loops play significant roles in the formation and stabilization of unimolecular G-quadruplex,[132, 177-179] loop-size dependency of our circularization reaction was thus examined during our investigation. Five new sequences were accordingly designed in our studies, in which the numbers of thymine nucleotide between guanine tracts are one (lp-1 in **Figure 4-10**), two (lp-2), three (lp-3), four (lp-4) and five (lp-5) respectively. As shown in **Figure 4-10**, our circularization reaction starting with lp-2 (Lane 4), lp-3 (Lane 6), lp-4 (Lane 8) and lp-5 (Lane 10 in **Figure 4-10**) proceeded in 31, 20, 72 and 60% yield respectively. Conversely there was absent of circularization product observable when the ligation reaction began with lp-1 (Lane 2). These results were consistent with the early suggestion[177] that at least two thymine nucleotides within the intervening loops are needed to sustain a stable fold-back structure of G-quadruplex.

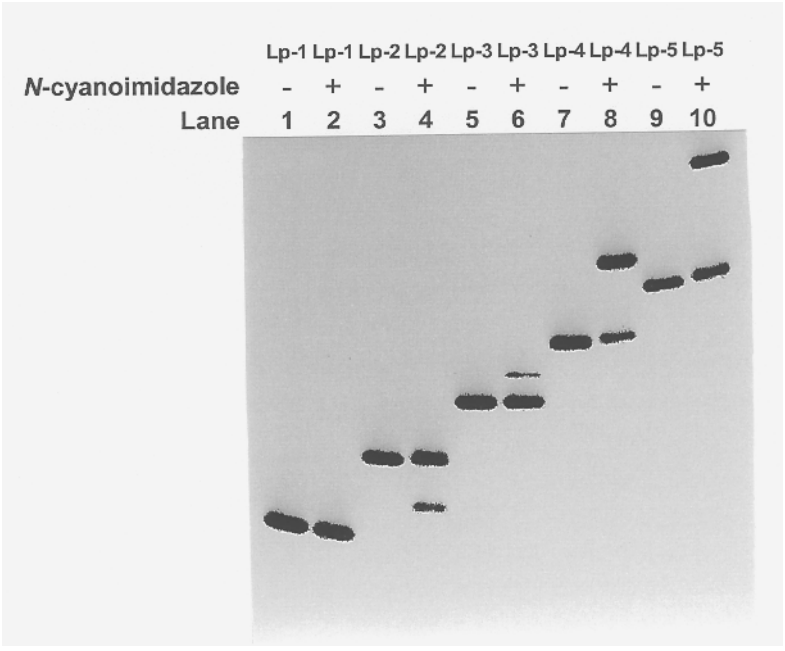


Figure 4-10. Effect of loop size on the circularization reaction. Reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for replacing **Sequence 1** with lp-1 (5' G_2**T**G_4**T**G_4**T**G_4**T**G_2 3'), lp-2 (5' G_2**T**$_2G_4$**T**$_2G_4$**T**$_2G_4$**T**$_2G_2$ 3'), lp-3 (5' G_2**T**$_3G_4$**T**$_3G_4$**T**$_3G_4$**T**$_3G_2$ 3') and lp-5 (5'

$G_2\mathbf{T_5}G_4\mathbf{T_5}G_4\mathbf{T_5}G_4\mathbf{T_5}G_2$ 3') respectively. Lane 1: lp-1 alone; Lane 2: reaction mixture of lp-1; Lane 3: lp-2 alone; Lane 4: reaction mixture of lp-2; Lane 5: lp-3 alone; Lane 6: reaction mixture of lp-3; Lane 7: **Sequence 1** alone; Lane 8: reaction mixture of **Sequence 1**; Lane 9: lp-5 alone; Lane 10: reaction mixture of lp-5.

4.2.5 Alkali-Ion Dependence of Our Circularization Course

Selective interaction with cations that fit well in the cavities formed by the stacking of guanine tetrads is a distinguishable characteristic of G-quadruplex from any other structural features of nucleic acids. In the alkali series, the order of ions preferred by G-quartet is $K^+ >> Na^+ > Rb^+ > Cs^+ > Li^+$.[175] With the intention of knowing whether alkali ion affects our circularization course, the efficiency of our circularization course in the presence of these ions was examined during our investigation. As shown in **Figure 4-11**, there was absent of circular product observable when lithium (Lane 2), sodium (Lane 3), rubidium (Lane 4) and cesium ions (Lane 5) were used to facilitate the formation of G-quadruplex structures. This took place most likely because these alkali ions might not be able to sustain a stable structure of G-quadruplex as potassium ion does under our standard condition for chemical ligation reactions.

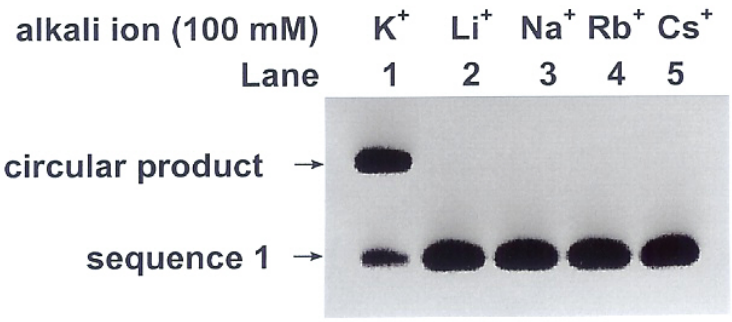


Figure 4-11. Effect of alkali ions on the circularization reaction. Lane 1. same as lane 9 in **Figure 4-5**. Lanes 2 to 5: Reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for replacing KCl with 100 mM of LiCl (lane 2), NaCl (lane 3), RbCl (lane 4) and CsCl (lane 5) respectively.

4.2.6 pH Dependence of the Designed Ligation Reactions

The effect of pH values on our circularization course was examined during our investigation in the range from 5.5 to 7.5. As shown in **Figure 4-12**, the efficiency of our ligation reactions decreased with the increase of pH values of the corresponding buffer solutions. These ligation reactions proceeded in 75, 78, 71, 49, 9 and 1% yield respectively at pH 5.5 (Lane 2), 6.0 (Lane 3), 6.5 (Lane 4), 7.0 (Lane 5) and 7.5 (Lane 6). Earlier CD spectroscopic analysis[177] revealed that variation of pH values of the corresponding buffer solutions ranged from 5 to 8 alone exhibited little effect on the stability of G-quadruplex. The relative low efficiency of our ligation reaction at higher pH values (Lanes 5 and 6) should be attributed to the fast decomposition of N-cyanoimidazole, the chemical agent required for the ligation reaction, under basic conditions.

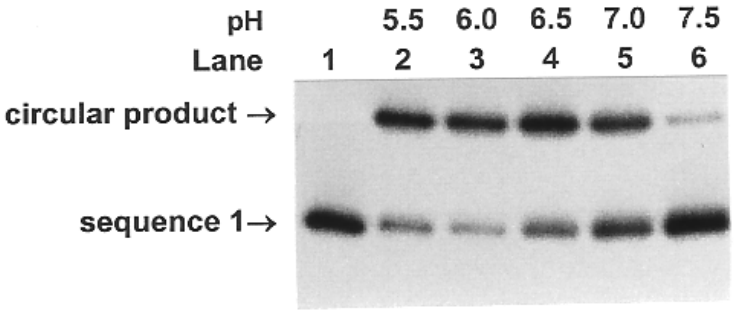


Figure 4-12. pH dependency of the circularization reaction. Reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for that pH of the buffer solutions were 5.5 (lane 2), 6.0 (lane 3), 6.5 (lane 4), 7.0 (lane 5), and 7.5 (lane 6) respectively; lane 1: **Sequence 1** alone.

4.2.7 Potassium Ion-Concentration Dependence of Our Ligation Reaction

Possession of cations by the structural feature of G-quartet as a part of its structure determines that the formation of G-quadruplex from its precursor of

unstructured sequences is an ion-concentration dependent process. Earlier capillary electrophoresis studies[128] revealed that G-quadruplex was not observable in the absence of cations while the amount of this tetraplex structure increased with the increase of the concentration of potassium ion. When the concentration of potassium chloride reached up to 100 mM,[128] single stranded form of DNA was converted to G-quadruplex structures completely. With the aim of finding out whether our designed unimolecular structure of G-quadruplex (e in **Figure 4-2**) is affected by the variation of ion concentration, our ligation reaction in the presence of different amount of potassium ion was accordingly examined during our investigation. As shown in **Figure 4-13**, the efficiency of the circularization reactions increased with the increase of potassium ion concentration (Lanes 2 to 8) and the ligation product in over 50% yield was obtained when potassium ion concentration was kept at 100 mM (Lane 2). These results were consistent with the early suggestion that there is equilibrium between the unstructured linear sequence and G-quadruplex structures in which a higher ion concentration favors the formation of structured DNA.[176]

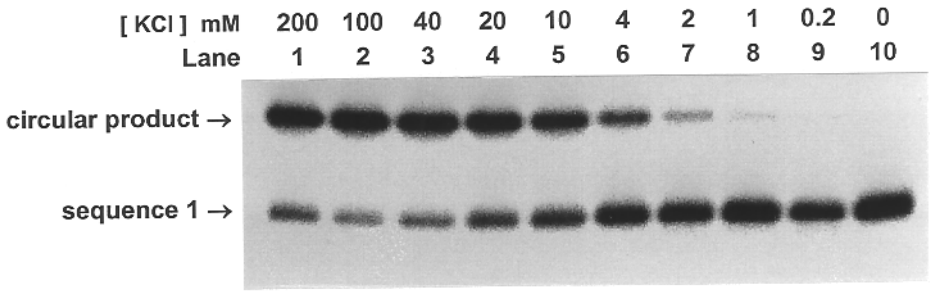


Figure 4-13. Effect of potassium-ion concentration on the circularization reaction. Ligation reactions were carried out in the same way as the one loaded in Lane 9 in **Figure 4-5** except for that the concentration of potassium chloride was kept at 200 mM (lane 1), 100 mM (lane 2), 40 mM (lane 3), 20 mM (lane 4), 10 mM (lane 5), 4 mM (lane 6), 2 mM (lane 7), 1 mM (lane 8), 0.2 mM (lane 9), 0 mM (lane 10) instead.

4.2.8 Verification of Formation of G-Quadruplex from Newly Synthesized Circular Oligonucleotides

With the aim of verifying that the structural feature of G-quadruplex could indeed be generated by the newly synthesized circular oligonucleotides, certain CD spectroscopic examinations were carried out on the corresponding ligation products. As shown in **Figure 4-14**, the CD spectrum of this circular oligonucleotide displayed a low amplitude peak near 275 nm in the absence of added salt. In the presence of potassium chloride, however, this circular oligonucleotide exhibited a spectrum characterized by positive maxima near 293 nm and a negative minimum near 265 nm, which are the typical features of formation of G-quadruplex from random conformations of oligonucleotides. The results shown in **Figure 4-14** are accordingly consistent with the suggestion that a guanine-rich circular oligonucleotide, like its linear counterpart, is capable of forming G-quadruplex in the presence of potassium chloride.

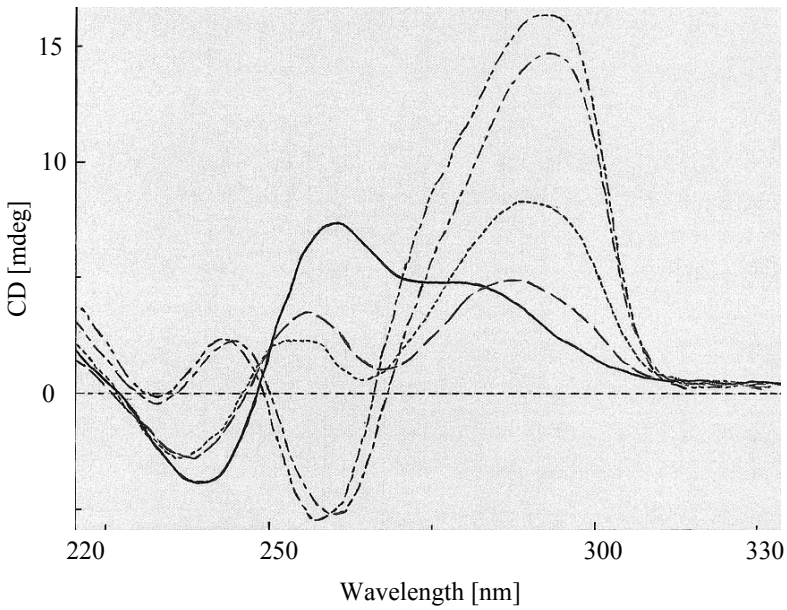


Figure 4-14. CD spectra of circular oligonucleotides of <GGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGG> (20 μM) in 10 mM Tris·HCl buffer (pH 7.0) containing 0 mM (——), 0.1 mM (— — —), 0.2 mM (– – – – – –), 2 mM (– — – —) and 20 mM (– – ——– –) KCl respectively.

Table 4-1. Sequences of oligonucleotides used in the current study.

Oligomer name	Nucleotide sequences
Sequence 1	5′-GGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGG-3′
Mis-1	5′-TGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGG-3′
Mis-2	5′-TGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGT-3′
Mis-3	5′-TTTTTTGGGGTTTTGGGGTTTTGGGGTTTTGT-3′
Mis-4	5′-TTTTTTGGGGTTTTGGGGTTTTGGGGTTTTTT-3′
Rec-1	5′-GTTTTGGGGTTTTGGGGTTTTGGGGTTTTGG-3′
Rec-2	5′-TTTTGGGGTTTTGGGGTTTTGGGGTTTTGG-3′
Rec-3	5′-TTTTGGGGTTTTGGGGTTTTGGGGTTTTG-3′
Rec-4	5′-TTTTGGGGTTTTGGGGTTTTGGGGTTTT-3′
Lp-1	5′-GGTGGGGTGGGGTGGGGTGG-3′
Lp-2	5′-GGTTGGGGTTGGGGTTGGGGTTGG-3′
Lp-3	5′-GGTTTGGGGTTTGGGGTTTGGGGTTTGG-3′
Lp-5	5′-GGTTTTTGGGGTTTTTGGGGTTTTTGGGGTTTTTGG-3′
Sequence 2	5′-GGTTTGGGTTTGGGTTTGGGTTTG-3′
Sequence 3	5′-GGTTAGGGGTTAGGGGTTAGGGGTTAGG-3′
G-0	5′-GGTTTTGGGGTTTTTTTTTTTGGGGTTTTGG-3′
G-1	5′-GGTTTTGGGGTTTTGTTTTTTTGGGGTTTTGG-3′
G-2	5′GGTTTTGGGGTTTTGGTTTTTTGGGGTTTTGG-3′
G-3	5′-GGTTTTGGGGTTTTGGGTTTTTGGGGTTTTGG-3′
45mer	5′-$T_3CGT_2CGTGCT_2GCT_3GCTCT_2G_3T_2AG_5T_2AG_2T_4$

4.3 Construction of Fluorescein-Tagged Circular Oligonucleotides

4.3.1 Design and Synthesis of Fluorescein-Tagged Circular Oligonucleotides

After verify the feasibility of constructing the unimolecular G-quadruplex without a fluorescein tag, a fluorescein molecule was labeled on **Sequence 1** for further examinations. **Figure 4-15** describes our approach of synthesis of circular G-quadruplex tagged with a fluorescein moiety. **Sequence 1** (**Figure 4-15**) is a linear 32-mer oligonucleotide containing a fluorescein moiety covalently linked to the carbon-5 of a thymine within the sequence. It was tentatively assumed in the early stage of these studies that this fluorescent moiety had little effect on the formation G-quadruplex since the covalently modified thymine would presumably appear in the loop region of G-quadruplex. Accordingly, a stock solution of **Sequence 1** (2μM) in a buffer solution containing 100 mM MES (pH 6.0) and 100 mM KCl was heated to 100 °C and kept at the same temperature for three minutes followed by cooling the mixture to 25 °C over a time period of two hours. A reaction mixture containing 1 μM of annealed **Sequence 1**, 100 mM MES (pH 6.0), 100 mM KCl, 10 mM $MnCl_2$ and 10 mM N-cyanoimidazole in a total volume of 20 μL were then prepared and incubated at 25 °C for 24 hours. The reaction was next terminated by the addition of loading buffer and the resultant ligation product was analyzed by polyacrylamide gel electrophoresis.

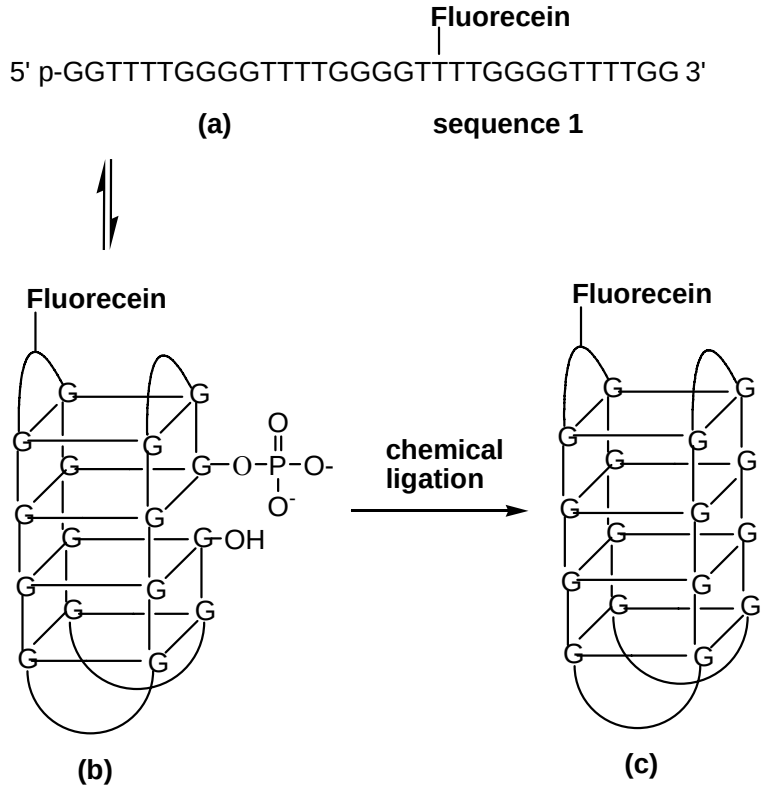


Figure 4-15. Schematic representation of our synthetic route toward fluorecein-labeled circular G-quadruplex.

As shown on the autoradiogram in **Figure 4-16**, a new product was formed from our designed ligation reaction with its mobility shift slower than that of **Sequence 1** (upper band of lane 5), which was presumably the desired fluorecein-labeled circular G-quadruplex (c in **Figure 4-15**). As controls, circularization reactions of a non-fluorecein- labeled oligonucleotide (5'pGGTTTTGGGGTTTTGGGGTTTTGGGGTTTTGG 3', **Sequence 2**) were also carried out (lanes 2) under the same condition as the one for the sample loaded into lanes 5 in **Figure 4-16**. As it can be seen on this autoradiogram, the circularization product of non-fluorecein-tagged oligonucleotide (Lane 2) possessed a faster rate of mobility shift than that of fluorecein-tagged oligonucleotide (Lane 5). The mobility shift difference between these two circular products can be attributed to the possession of an extra fluorecein moiety within the fluorescein-tagged circular

product (c in **Figure 4-15**) in our designed reaction.

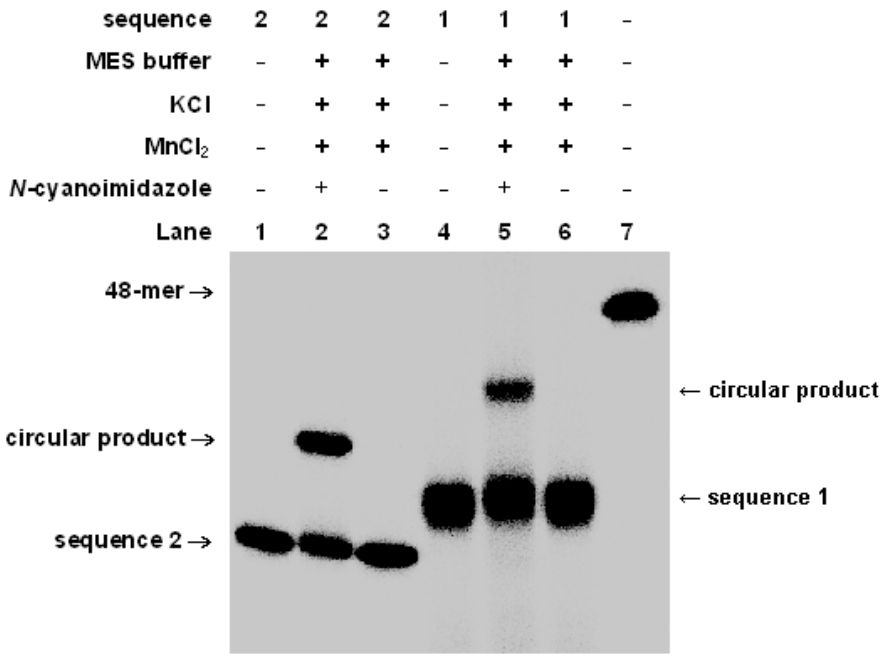


Figure 4-16. Electrophoretic analysis of fluorecein-labeled circular G-quadruplex. Lane 1: **Sequence 2** alone. Lane 2: Same reaction as the one loaded in Lane 5 except for replacing **Sequence 1** with **Sequence 2**. Lane 3: same reaction as the loaded in Lane 2 except for the absence of N-cyanoimidazole. Lane 4: **Sequence 1** alone. Lane 5: a reaction mixture containing 1 μM annealed **Sequence 1**, 100 mM MES (pH 6.0), 100 mM KCl, 10 mM $MnCl_2$ and 10 mM N-cyanoimidazole in a total volume of 20 μL was incubated at 25 °C for 24 hours. Lane 6: same reaction as the loaded in Lane 5 except for the absence of N-cyanoimidazole. Lane 7: the 48-mer of 5'$T_3CGT_2CGTGCT_2GCT_3GCT_3C_2T_2G_3T_2AG_5T_2AG_2T4$ 3' as a molecular weight marker.

4.3.2 Structural Verification of Fluorescein-Tagged Circular Oligonucleotides

With the aim of verifying that circularization of fluorescein-tagged circular oligonucleotides indeed took place, the newly formed product (upper band in Lane 5 of **Figure 4-16**) was purified and further digested with exonuclease VII (an exodeoxyribonuclease that hydrolyzes linear DNA from both 3' and 5' termini). Due to the absence of open ends in its structure, fluorecein-labeled circular oligonucleotide should in theory resist the degradation by this exonuclease. [180, 181] As shown in **Figure**

4-17, the newly formed product in the circularization reaction completely resisted the hydrolysis by exonuclease VII (Lane 4) as expected. In contrast, a linear sequence (**Sequence 2**) was hydrolyzed to completion by exonuclease VII (lane 2) under the same reaction condition.

In addition to this test of degradation by exodeoxyribonuclease, hydrolysis of the fluorecien-labeled circular G-quadruplex using DNAse I was carried out. DNAse I is an endonuclease hydrolyzing both single and duplex forms of DNA in a randomized pattern. As shown in **Figure 4-18**, a major band was observed upon partial hydrolysis of our fluorescein-labeled circular product by the endonuclease. This degradation product displayed the same rate of mobility shift as that of its linear precursor (Lane 3), which was the characteristic pattern of randomized partial hydrolysis of circular oligonucleotides by enzyme or chemicals.[182, 183] The results of both exonuclease and endonuclease hydrolysis in these studies are consistent with the suggestion that the newly formed product in our designed reaction is indeed circular in its backbone structure.

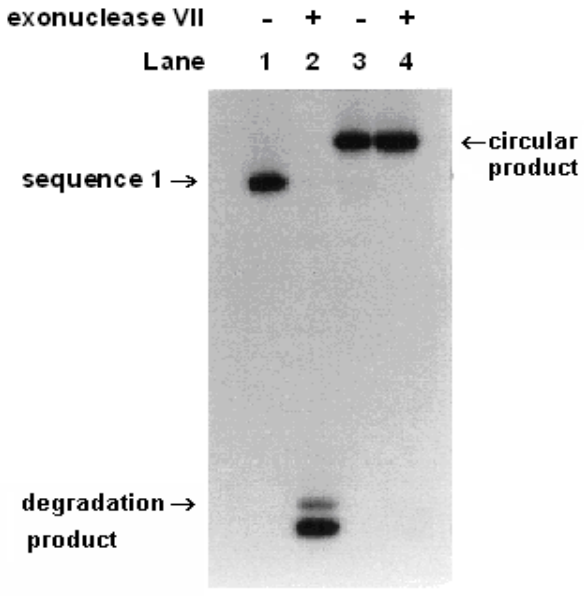


Figure 4-17. Hydrolysis of fluorecein-labeled circular products by exonuclease VII.

Lane 1: **Sequence 1** alone. Lane 2: **Sequence 1** after treatment with 20 units of exonuclease VII at 37 °C for 2 h. Lane 3: our designed fluorecein-labeled circular G-quadruplex alone. Lane 4: our designed fluorecein-labeled circular G-quadruplex after treatment with 20 units of exonuclease VII at 37 °C for 2 h.

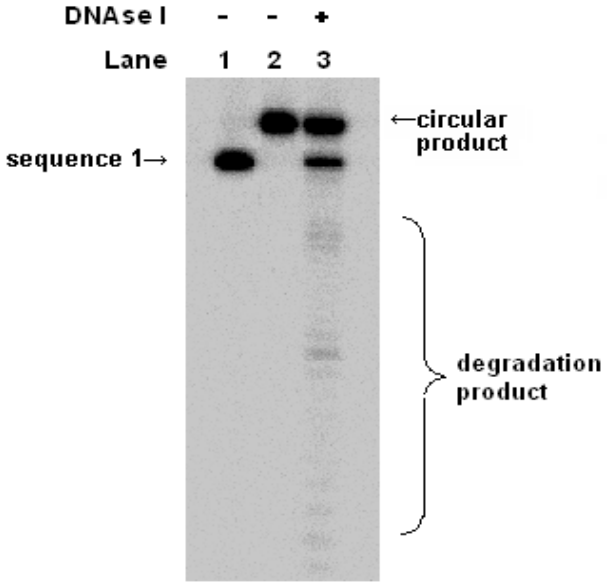


Figure 4-18. Partial hydrolysis of the fluorecein-labeled circular products by DNAse I. Lane 1: **Sequence 1** alone. Lane 2: our designed fluorecein-labeled circular G-quadruplex alone. Lane 3: our designed fluorecein-labeled circular G-quadruplex after treatment with 10 units of DNAse I at 25 °C for 30 min.

4.3.3 Fluorescence Measurement of Fluorescein-Tagged Circular G-Quadruplex

It was well established in the past that when fluorescein-labeled oligonucleotides were under fluorescence spectroscopic examination, the corresponding emission is readily detectable.[184] In order to verify that the molecular moiety of fluorescein was indeed present in our designed circular G-quadruplex, fluorescence spectroscopic measurements were carried out on the corresponding samples of oligonucleotides. As shown in **Figure 4-19**, our designed circular G-quadruplex exhibited the emission maximum at 520 nm when the wavelength of

excitation was set at 495nm while a non-fluorescence-tagged oligonucleotide (**Sequence 2**) displayed no emission in the corresponding range of wavelength. These results are the indication that fluorescein moiety is a part of the molecular entity of our designed circular G-quadruplex.

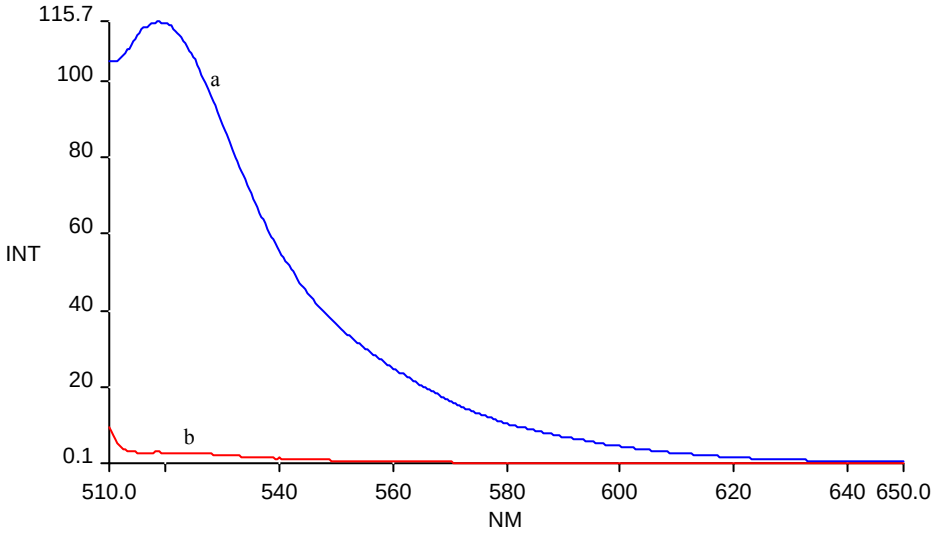


Figure 4-19. Fluorescence emission spectra of fluorescein-labeled circular G-quadruplex (a) and non-fluorescein-labeled linear oligonucleotide, **Sequence 2** (b). Samples in the measurements contained 1 μM oligonucleotides and 10 mM Tris buffer (pH 7.0) and were examined at 20 °C with excitation set at 495 nm

4.4 Conclusions

This work demonstrates that certain unimolecularly circular G-quadruplexes can indeed be attained on the template basis of G-quadruplex through chemical ligation reactions. These circularization courses are highly effective and sequence specific under our standard reaction conditions. The circular structure of G-quadruplex possessing two, three, four or five thymine nucleotides in its intervening loops is readily achievable while the linear sequences possessing a single thymine nucleotide between its guanine tracts is not a proper precursor for our circularization

reaction.

In addition, it was demonstrated in our studies that the presence of potassium ion is the prerequisite for the synthesis of circular G-quadruplex while such circular sturcutres could not be obtained in the presence of other alkali ions such as Na^+ and Li^+. Moreover, our further examinations showed that this circularization course can be affected by other factors such as pH values of the corresponding buffer solutions as well as the concentration of potassium ions.

On the basis of our successful synthesis of circular G-quadruplex, fluorescein-labeled circular G-quadruplex was also accomplished during our studies. This newly obtained fluorescein-tagged circular product resisted the hydrolysis by an exonuclease as anticipated. In addition, fluorescence spectroscopic analysis revealed that fluorecein moiety was indeed associated with the designed G-quadruplex.

It is anticipated that the unimolecularly circular G-quadruplex with a fluorescein tag could act as useful molecular probes for identifying new types of enzymes and other functional proteins correlated with G-quadruplex *in vivo*. In addition, it is our hope that the newly designed fluorescein-tagged G-quadruplex could be utilized for tracing the location and distribution of G-quadruplex-binding proteins in living cells through detection of the signal of fluorecein covalently linked to circular G-quadruplex.

Chapter 5

Development of New Oligonucleotides-Based Topoisomerase I Inhibitors

5.1 Background and Aims

5.1.1 DNA Topoisomerases

In eukaryotic cells, double helical DNA is supercoiled and condensed into chromosomes by wrapping tightly around histones.[185, 186] For a number of essential cellular processes, the information stored in DNA must be accessed, thus the two strands of the helix have to be separated. In addition, during transcription, the translocating RNA polymerase generates supercoiling tension in the DNA which also needs to be relaxed. Besides, replications of two complementary DNA strand, or recombination between DNA duplex can generate chromosomal knots or catenanes.[187] Failure to solve these topological problems would cause chromosomal breakage or even severer consequence for cell viability.

DNA Topoisomerases, the marvelous molecular machines designed by nature, happen to be such an essential class of enzymes that can overcome the above topological problems generated by multiple cellular processes.[188] These enzymes catalyze topological rearrangements of DNA through sequential single-stranded breakage, strand free rotation and rejoining of phosphodiester backbone of DNA,

which removes knots and catenanes generated by metabolic processes *in vivo*.[189]

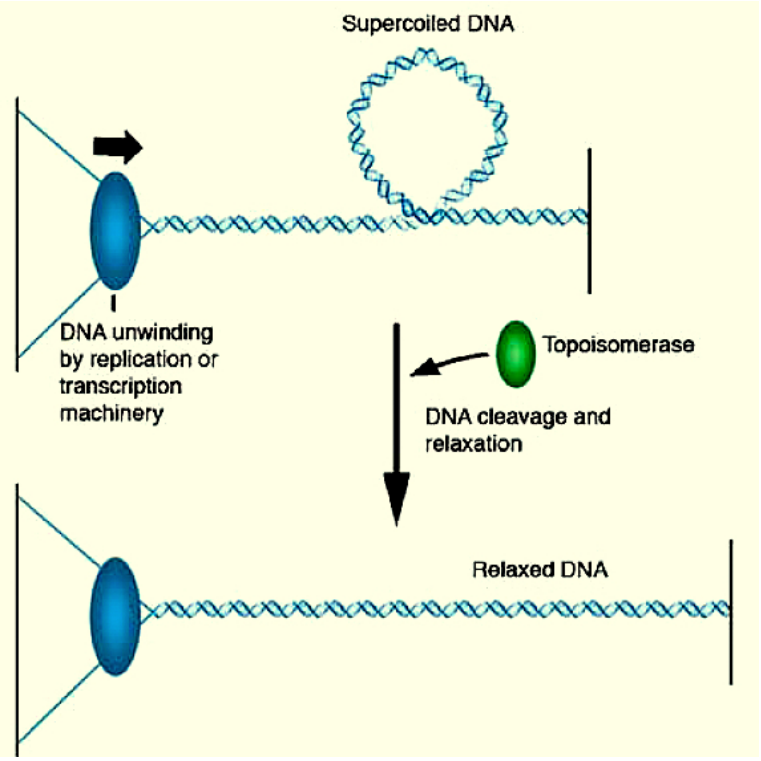


Figure 5-1. Superhelical tension generated by DNA unwinding and resolved by DNA topoisomerases.[1]

Topoisomerases can be grouped into two types, and each of them has a unique structure and functionality. The enzymes that cleave only one strand of the duplex DNA are defined as type I topoisomerases, which can be further classified as either type IA subfamily members in which the protein links to a 5' phosphate or type IB subfamily members in which the protein attaches to a 3' phosphate. Topoisomerases that cleave both strands to generate a staggered double-strand break belong to the type II subfamilies.[187]

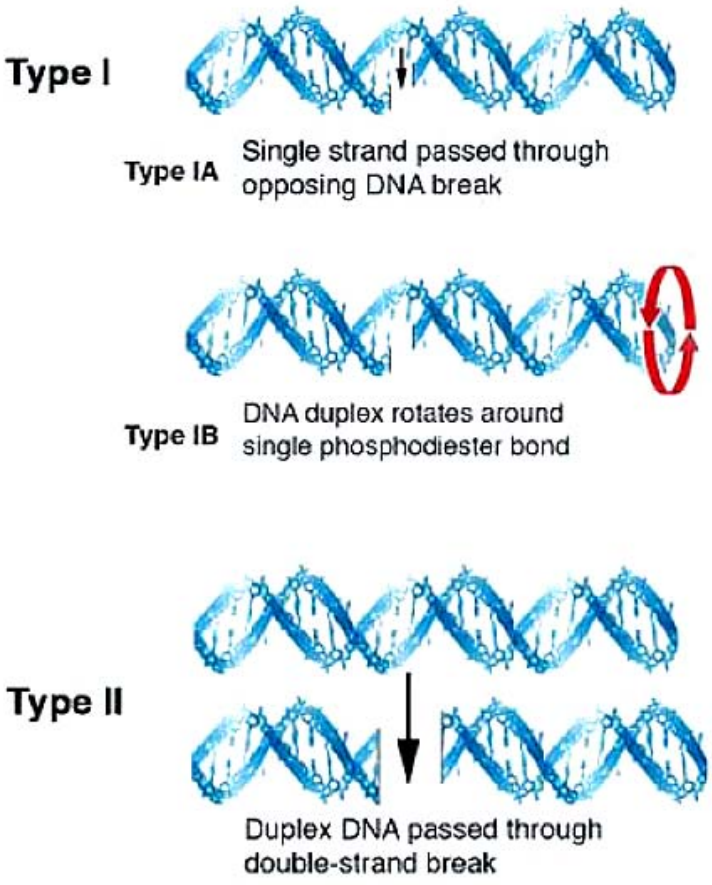


Figure 5-2. Type I and Type II DNA topoisomerases[185]

5.1.2 Mode of Action of DNA Topoisomerase I

The DNA relaxation activity of DNA topoisomerase I does not require ATP or Mg^{2+}. They bind only to double stranded DNA and change the linking number of DNA by one in each cycle. The relaxation of supercoiled DNA by topoisomerase I involves these essential steps: DNA binding (a) , nucleophilic attack by the active site tyrosine leads to cleavage of the DNA backbone (b), single strand rotation (c), relegation of phosphate backbone to restore the continuity of DNA (d) and finally released from DNA (e and f). Initially, DNA topoisomerase I binds to double-stranded DNA and covers a region of about 20 bp. It then cleaves one of the double helical DNA strands by launching a nucleophilic attack on the 3' end scissile phosphoryl of the DNA backbone through a reversible trans- esterification reaction. The cleavage

action results in a covalent attachment of the 3'-phosphate onto the protein and a free 5'-hydroxyl end. Besides, the cleavage leads to the formation of a transient covalent topoisomerase I - DNA intermediate. Subsequently, the tension within the supercoiled DNA drives a controlled rotation of the free 5' end to achieve a low energy conformation and hence relaxed form of DNA. The enzyme subsequently religates the 5'-hydroxyl termini of the nicked strand and dissociates itself from the relaxed DNA.[186-189]

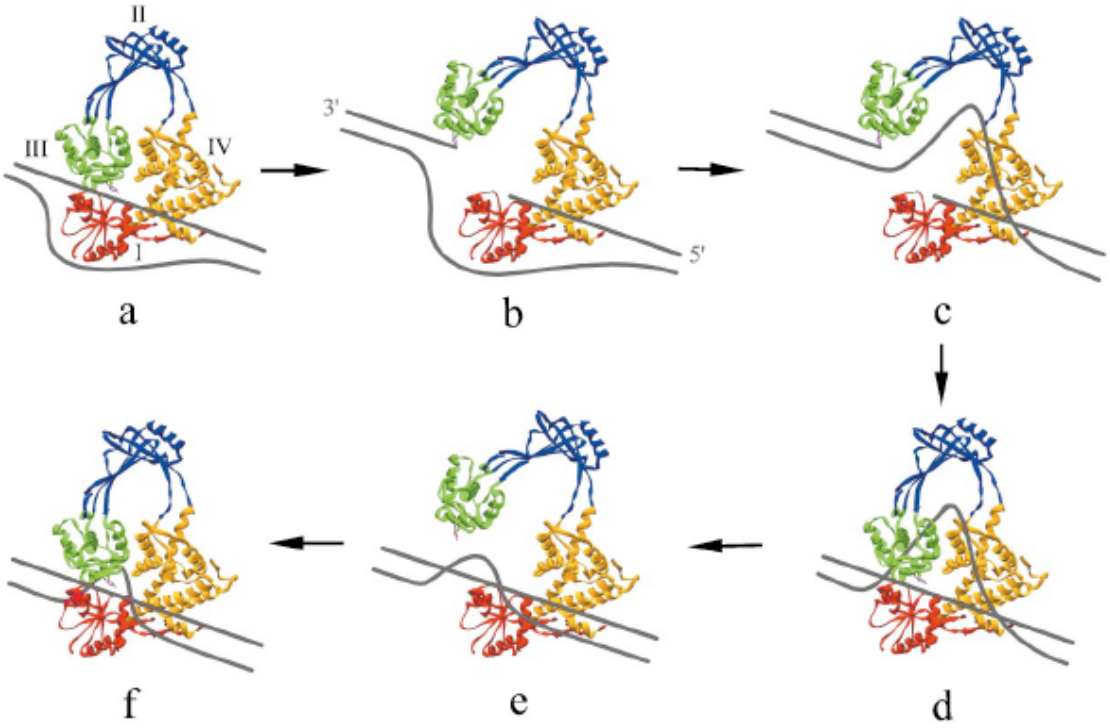


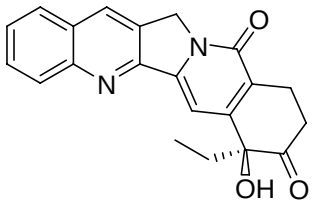
Figure 5-3. Mode of action of DNA topoisomerase I [186]

The above mentioned pathway of action of topoisomerase I is known as the swivel mechanism, which is specific to topoisomerase IB, the class of enzyme that human topoisomerase I belongs to. Other classes of topoisomerases include topoisomerase IA, which employs the strand passage mechanism of relaxing supercoil, while topoisomerase IIα and topoisomerase IIβ cleaves two strands of DNA in its mode of action.

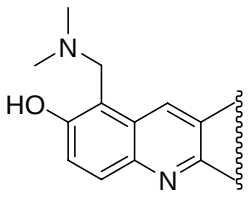
5.1.3 Topoisomerase I inhibitors

Growing interest in toposiomerases not only attributes to their ability of managing DNA topology, but also to their importance in pharmacology and clinical medicine. Topoisomerase I has been found to be highly abundant in fast proliferating tumor cells than in normal tissue[188, 189] and identified as a molecular target of several anticancer agents. Camptothecin, irinotecan and topotecan are the examples of topoisomerase I inhibitors that are currently used in the clinical treatment of cancers (**Figure 5-4**).[190]

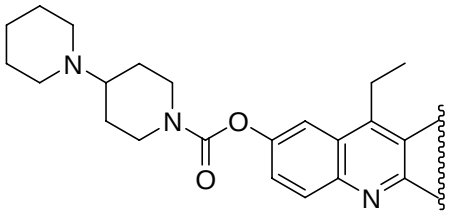
Camptothecin was the first specific topoisomerase I inhibitor discovered and was first identified as a drug with potential antitumor activity in 1966.[190] It was originally isolated from the Chinese tree Camptotheca acuminate. Irinotecan and topotecan are the derivatives of camptothecin.[190] In the process of actions, these anticancer agents bind to a transient topoisomerase I-DNA covalent complex and inhibit the resealing of a single-strand nick that the enzyme creates to relieve superhelical tension in duplex DNA, functioning as poisons of the enzyme. Although these drugs exhibit anti-tumor activity, they have some undesirable side effects. For example, irinotecan causes symptoms such as cholinergic-like syndrome, delayed onset diarrhea, neutropenia, nausea and vomiting, fatigue and alopecia.[190]



Camptothecin



Topotecan (hycamptamine)



CPT-11(irinotecan)

Figure 5-4. Chemical structures of Camptothecin, Topotecan and irinotecan.

Besides these natural products and synthetic organic compounds identified as inhibitors of DNA topoisomerase I, certain oligonucleotides-based inhibitors were designed earlier in our laboratory that displayed high inhibitory potency on the activity of this DNA relaxing enzyme.[191] In order to explore new oligonucleotides for topoisomerase I inhibitors, certain particular structural assemblies of oligonucleotides were designed and synthesized during my graduate studies. The corresponding results are discussed in the following sections of this chapter.

5.2 Construction of C3-spacer-Containing Circular Oligonucleotides as Topoisomerase I Inhibitors

5.2.1 General Design Strategy

Due to their possession of mismatched base pairs,[198] the early designed oligonucleotide inhibitors reported in our previous studies could be vulnerable to the degradation by some DNA repair enzymes. With the aim of developing oligonucleotide inhibitors that could resist the hydrolysis by mismatch and nick-resolving DNA repair enzymes, we have very recently introduced C3-spacer (-CH2-CH2-CH2-) modifications[199] into the corresponding oligonucleotide structures.

Figure 5-5 shows a dumbbell-shaped circular oligonucleotide[200] (**Oligonucleotide 1**) designed during our recent studies, which possesses two

"extraordinarily thermostable hairpins"[201] at both termini of its duplex structure. It is anticipated that introduction of these two hairpins to **Oligonucleotide 1** could lead to an increase of thermal stability of the dumbbell-shaped structure and prevent itself from hydrolysis by exonucleases.[201] In addition, it has been well established in the past that during the strand scission and religation process, Topo I forms transitorily a covalent bond with the 3' end of DNA fragment and holds the 5' end of the second fragment through physical interaction. [192-197, 202]

Taking advantage of the fragile interacting fashion between Topo I and the resultant 5' end of DNA, a C3-spacer was introduced into the dumbbell-shaped structure near the Topo I cutting site of **Oligonucleotide 1** in order to generate an irreversible inhibition to the nuclear enzyme. It was our expectation that once Topo I would bind to **Oligonucleotide 1** and further cause a strand scission at its cutting site[203] (**Figure 5-6**), the religation reaction between the cut fragments might not readily take place. This could happen because the presence of C3-spacer in the dumbbell-shaped structure might lead to a free dissociation of the corresponding trinucleotide (5' GGA, see **Figure 5-6**) from Topo I and its complementary stretch, thus leading to chemically irreversible damage to the enzyme.[204, 205] In addition, it was anticipated that **Oligonucleotide 1** could resist the hydrolysis by some mismatch- and nick-resolving DNA repair enzymes since introduction of C3-space would not lead to any base mismatches or open 5' and 3' ends in the dumbbell-shaped structures.

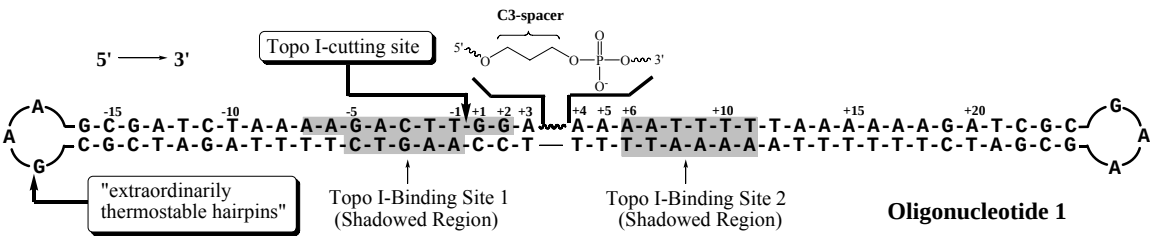


Figure 5-5. Schematic representation of a C3-spacer-containing dumbbell-shaped oligonucleotide designed in our studies. [192-197]

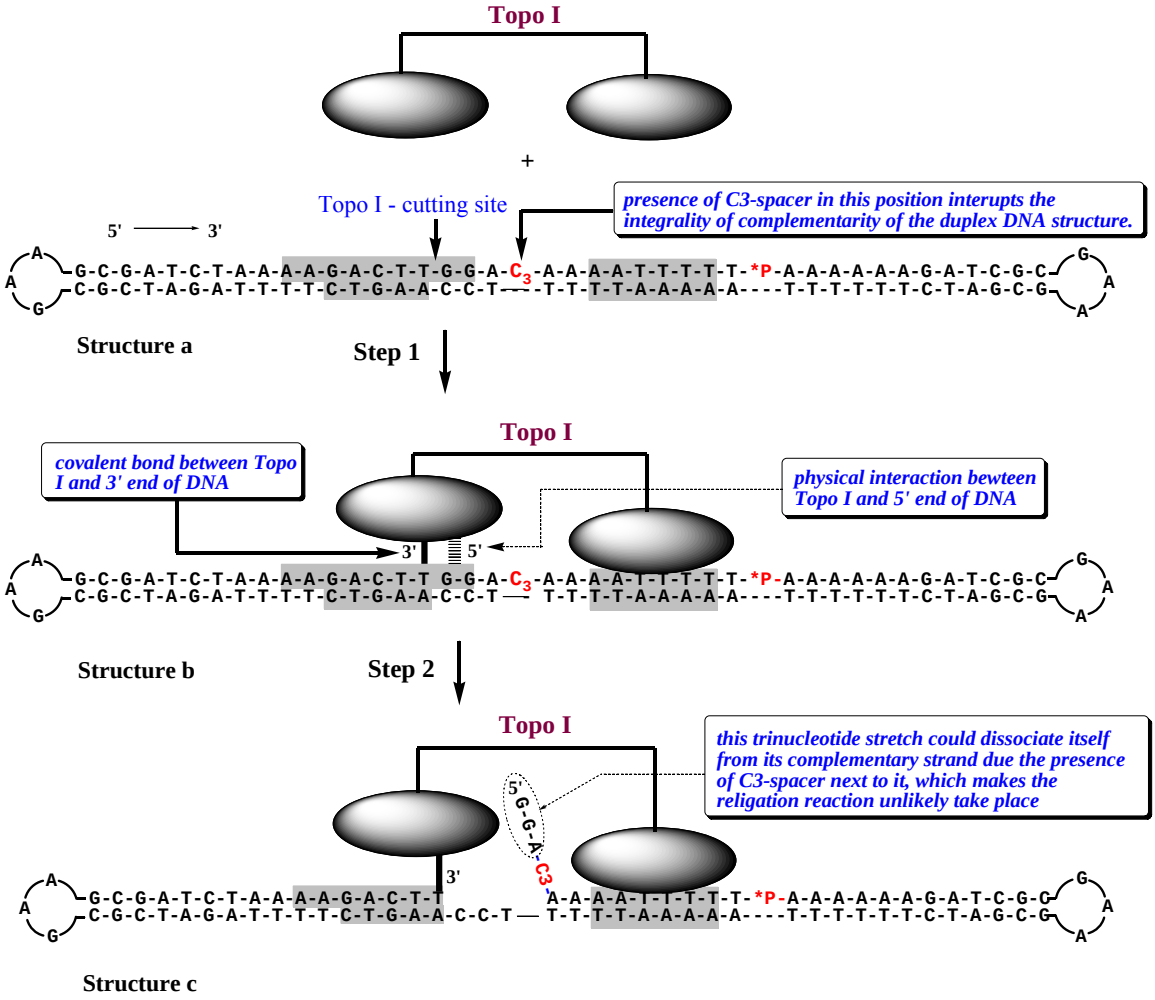


Figure 5-6. Diagrammatic illustration of anticipated inhibitory mechanisms of a C3-spacer-containing oligonucleotide (**Oligonucleotide 1**) on the activities of human topoisomerase I in our studies. When **Oligonucleotide 1** is incubated with Topo I (Step 1), this nuclear enzyme is expected to bind to the oligonucleotide and subsequently form a covalent bond with the newly generated 3' end of DNA and to hold the 5' end through physical interaction (Structure b). The presence of C3-spacer near the newly generated 5' end will interrupt the integrity of complementarity of the local duplex structure, which could consequently allow the trinucleotide of 5' G-G-A- to dissociate freely from its complementary tract (Structure c). This free dissociation of 5' G-G-A- from its complementary sequence could make the religation reaction impossible, thus leading an irreversible damage to the enzyme.

5.2.2 Synthesis and Characterization of C3-Spacer-Containing Circular Oligonucleotides

Synthesis of **Oligonucleotide 1** was subsequently carried out in this study through enzymatic ligation reaction (**Figure 5-7**). The 5' phosphorylated linear

precursor of 86-mer oligonucleotide containing a C3-spacer modification was provided by Sigma-Proligo. By using its own duplex backbone as template, the two termini were ligated at 15 °C in the presence of T4 DNA ligase. The results were analyzed via 15% denaturing PAGE. As shown in **Figure 5-8**, a band with faster mobility shift appeared after 2 hours and the circulization was almost completed in 12 hours. Circularity of the backbone of the formed **Oligonucleotide 1** was further confirmed via hydrolysis using T7 exonuclease (**Figure 5-9**).

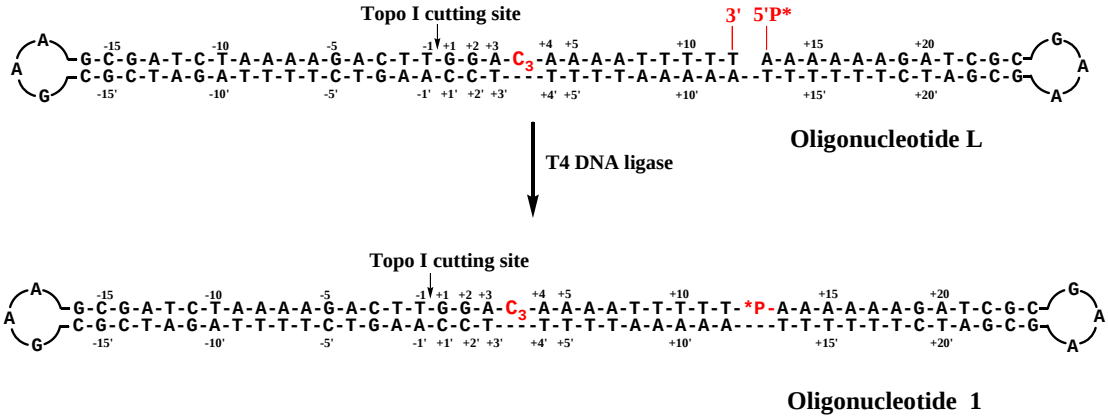


Figure 5-7. Illustration of the ligation reaction of **Oligonucleotide 1**

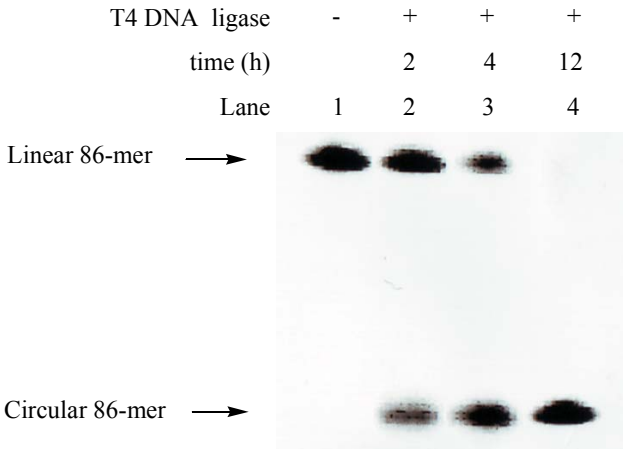


Figure 5-8. Polyacrylamide gel electrophoretic analysis of formation of **Oligonucleotide 1**. 86-mer linear precursor (Olignucleotide L in **Table 5-2**) with free 5'-hydroxyl terminus was labeled at its 5'-end with [γ-^{32}P]ATP (Ge-Healthcare) catalyzed by T4 polynucleotide kinase (New England Biolabs). A mixture containing 50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 1mM ATP, 10 mM dithiothreitol, 25 lg/ml BSA, and 10 μM ^{32}P-labeled 86-mer linear precursor was then kept at 95°C for 5 min and further allowed to cool down to 25°C over 2 h. After the reaction mixture was incubated at 4 °C for another 4 h, 10 units of T4 DNA ligase was added, which was

further incubated at 15 °C for different time periods. Lane 1: linear precursor (**Oligonucleotide L**) alone, lanes 2–4: ligation reaction of **Oligonucleotide L** catalyzed by T4 DNA ligase lasting for 2, 4, and 12 h, respectively.

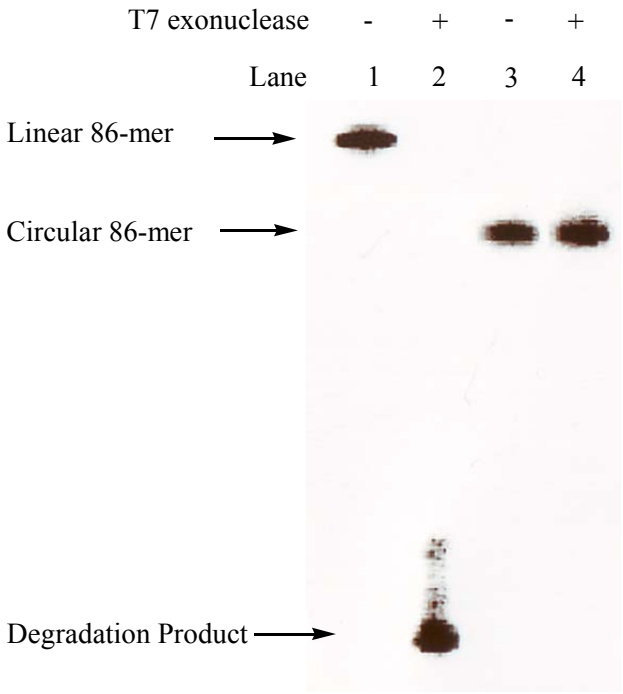


Figure 5-9. Polyacrylamide gel electrophoretic analysis of circularity of **Oligonucleotide 1** in its backbone. The ^{32}P-labeled reaction products in the lower band of Lane 4 (presumably **Oligonucleotide 1**) in **Figure 5-8** were cut off and purified. Solutions containing 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM Magnesium Acetate, 1 mM Dithiothreitol, 5 units of T7 exonuclease and ^{32}P-labeled oligonucleotides were incubated next at 25 °C for 2 h. Lane 1: ^{32}P labeled 5' phosphorylated 86-mer linear precursor (**Oligonucleotide L**) alone, Lane 2: ^{32}P labeled 5' phosphorylated 86-mer **Oligonucleotide L** in the presence of T7 exonuclease, Lane 3: **Oligonucleotide 1** (purified from the lower band in lane 4 in **Figure 5-7**) alone, Lane 4: **Oligonucleotide 1**(purified from the lower band in lane 4 in **Figure 5-7**) in the presence of T7 exonuclease.

5.2.3 Inhibitory Studies of the Newly Designed Oligonucleotides on the Activity of Topoisomerase I Inhibitors

The inhibitory effect of **Oligonucleotide 1** on the activity of human Topo I was examined through determining the efficiency of relaxation reaction of pBR322 catalyzed by this DNA relaxing enzyme in the presence of our newly designed dumbbell-shaped oligonucleotides. As shown in **Figure 5-10**, the ratio of relaxed

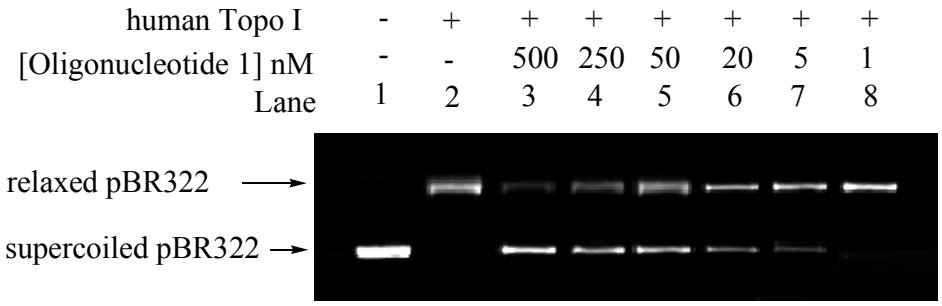
forms to non-relaxed forms of pBR322 decreased with the increase of concentration of **Oligonucleotide 1** (**Figure 5-10b**) and the obtained IC_{50} value for **Oligonucleotide 1**[206] was 36.6 nM (**Figure 5-11**) under our reaction conditions. As a control experiment, the inhibitory effect of **Oligonucleotide 2** was also examined in our studies, which possesses the same sequence as that of **Oligonucleotide 1** except for the absence of a C3-spacer in its duplex structure. As shown in **Figure 5-10c**, this non-C3-spacer-containing oligonucleotide displayed much low inhibitory efficiency and gave rise to an IC_{50} value as high as ~1.2 μM. The observed difference in inhibitory efficiency between **Oligonucleotide 1** and **Oligonucleotide 2** could be taken as an indication that C3-spacer modification indeed plays a crucial role in the inhibitory action on the activity of Topo I as originally designed (**Figure 5-6**).

a) supercoiled pBR322 + human Topo I

↓ newly designed oligonucleotide inhibitors

reduction of generation of relaxed pBR322

b)



c)

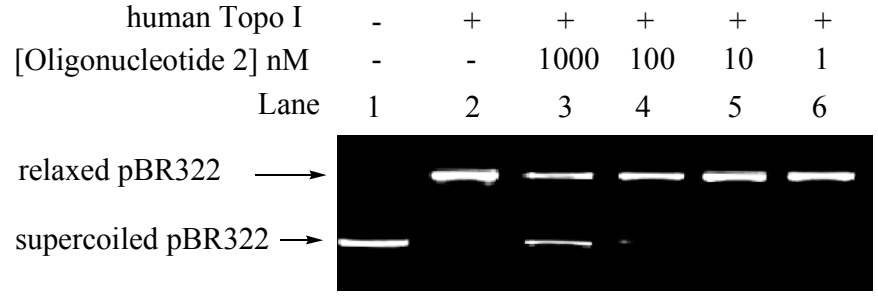


Figure 5-10. Agarose gel electrophoretic analysis of inhibitory effect of **Oligonucleotide 1** (b) and **Oligonucleotide 2** (c) on the activities of human topoisomerase I. **Oligonucleotide 1** or **2** and 1 unit human topoisomerase 1 (Topogen, 2005H-2) were incubated in a buffer of 10 mM Tris-HCl (pH 7.9), 150 mM NaCl, 0.1% bovine serum albumin (BSA), 0.1 mM spermidine, and 5% glycerol in a total volume of 20 μL at 25 °C for 5 mins. 400 ng pBR322 was following added and the new mixtures were further allowed to react at 37 °C for 20 mins. The samples were analyzed using agarose gel (1.0%, w/v) eletrophoresis (100V for 1h) followed by ethidium bromide staining. The DNA bands were visualized using Gel Documentatoin system (G:Box HR, Syngnene, Cambridge, UK) equipped with Gene Tools Software. b) Concentrations of **Oligonucleotide 1** in lane 2-8 were 0 nM, 500 nM, 250 nM, 50 nM, 25 nM, 5 nM, 1 nM respectively. Lane 1 is pBR 322 alone. c) Concentrations of **Oligonucleotide 2** in lane 2-6 were 0 nM, 1 nM, 10 nM, 100 nM, 1000 nM respectively. Lane 1 is pBR 322 alone.

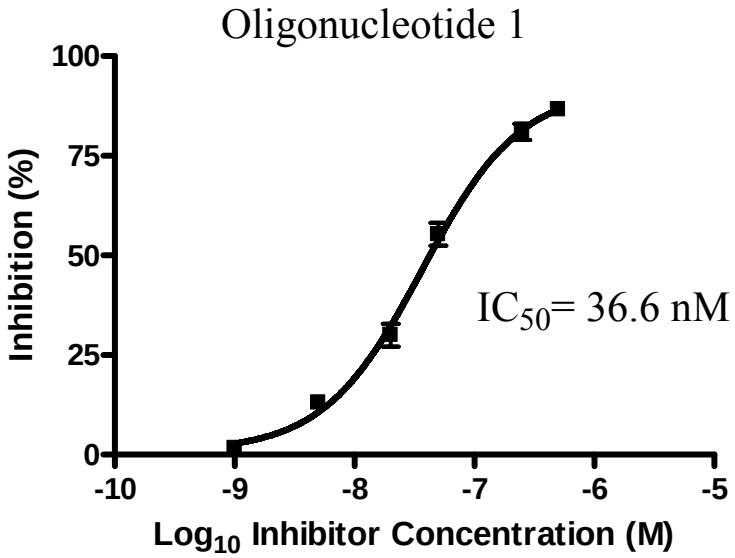


Figure 5-11. Correlations between concentration of **Oligonucleotide 1** and percent inhibition on topoisomerase I activity. Percentage of relaxation was defined as the ratio of band density of relaxed DNA over those of relaxed DNA plus supercoiled DNA [relaxed DNA/ (relaxed DNA + supercoil DNA)] while (100%- percentage of relaxation) was taken as the percent inhibition of topoisomerase I activity by oligonucleotides. The DNA bands were analyzed using Gel Documentation System (G:Box HR, Syngnene, Cambridge, UK) equipped with Gene Tools Software. IC_{50} value is calculated by GraphPad Prism 4 software.

5.2.4 Confirmation of the Existence of Topoi-DNA Covalent Conjugate

In order to verify that some covalent conjugates could be indeed formed between Topo I and **Oligonucleotide 1** in the inhibitory process shown in **Figure 5-10b** as designed, ^{32}P-labeled **Oligonucleotide 1** was incubated with Topo I during our

investigations and the mixture was analyzed through denaturing polyacrylamide gel electrophoresis. As shown in **Figure 5-12**, with the increase of amount of Topo I from 4 units to 16 units, new slow moving bands (Band 2) were observed (lane 2 to lane 7). Generation of the new bands (Band 2) in these studies could be considered as the sign that covalent conjugates between Topo I and **Oligonucleotide 1** were formed under our reaction conditions as anticipated.

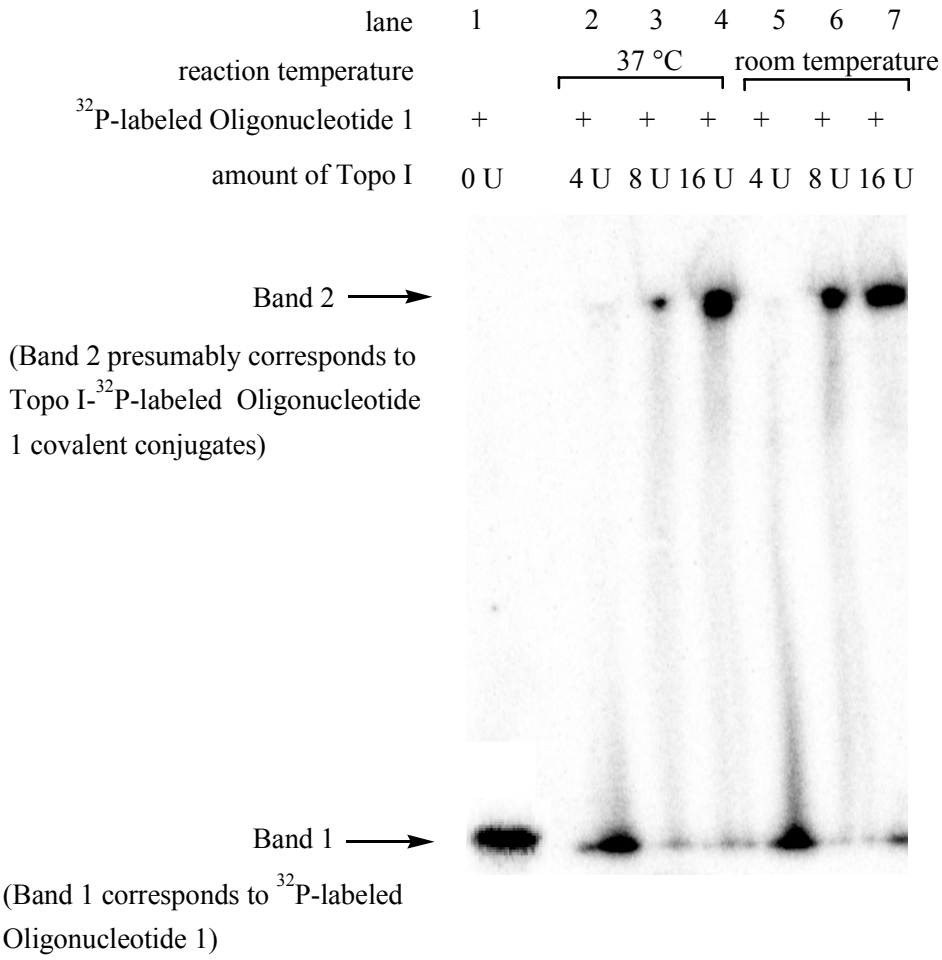


Figure 5-12. Denaturing polyacrylamide gel electrophoretic confirmation of formation of Topo I-**Oligonucleotide 1** covalent conjugates (see Structure c in **Figure 5-6**). Solutions (20 μL) containing ^{32}P-labeled **Oligonucleotide 1**, Topo I, 1X Topo I reaction buffer (10 mM Tris-Cl at pH 7.9, 150 mM NaCl, 0.1 % BSA, 0.1 mM Spermidine, 5 % glycerol), 10 mM $MgCl_2$, 50 mM KCl were incubated at 37 °C or room temperature for 1 h, which were further loaded on denaturing polyacrylamide gel [6% PAGE (29:1) containing 7 M urea]. Lane 1: ^{32}P-labeled **Oligonucleotide 1** alone; lane 2 to lane 4: 4 U, 8 U and 16 U Topo I respectively (reaction carried out at 37 °C), lane 5 to lane 7: 4 U, 8 U and 10 U Topo I respectively (reaction carried out at room temperature).

5.2.5 Examination of Resistance of Oligonucleotide 1 against Repair Enzyme

To maintain the integrity of its genome and normal functioning, cell adopts intricate enzymatic systems for detecting and resolving DNA damages generated by environmental factors and normal metabolic processes.[207] A number of DNA repair enzymes such as those that recognize mismatch base pairs and nicked sites have been identified in the past,. One of our major intentions in placing a C3-spacer into **Oligonucleotide 1** is to prevent its degradation by DNA repair proteins for taking place as neither mismatch base pairs nor nicked sites will be generated upon introduction of this modification.[208] As model studies, T7 endonuclease I, a DNA repair protein that removes mismatch base pairs and nicked sites from duplex DNA sequences, was selected during our investigations. This DNA repair protein was accordingly allowed to react with a mismatch-containing dumbbell-shaped structure (**Oligonucleotide 3**),[198, 208] which was consequently hydrolyzed into low molecular weight fragments within 1 hour (lane 4). However, when **Oligonucleotide 1** (C3-spacer-containing dumbbell-shaped structure) and **Oligonucleotide 2** (regular dumbbell-shaped structure without modification) were incubated with T7 endonuclease I, no degradation product was generated (lane 6 and lane 2 in **Figure 5-13**). These observations indicate that our newly designed C3-spacer-containing oligonucleotides could indeed resist DNA repair protein as expected.

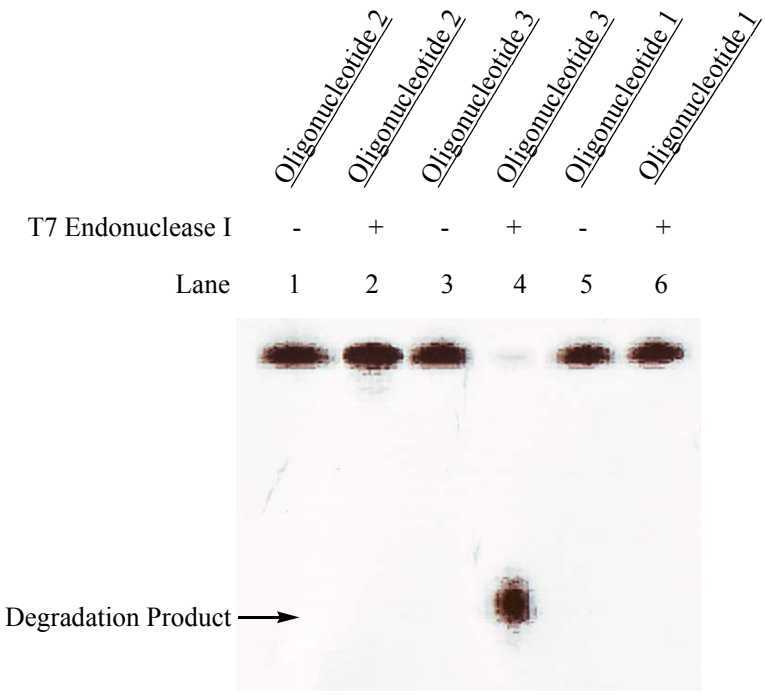


Figure 5-13. Polyacrylamide gel electrophoretic analysis of hydrolytic products of **Oligonucleotide 1, 2** and **3** generated by T7 endonuclease I. Solutions containing 50 mM NaCl, 10 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 1 mM Dithiothreitol, 10 U of T7 endonuclease I and ^{32}P-labeled oligonucleotides (**Oligonucleotide 1, 2** or **3**) were incubated at 37 °C for 1 h. The hydrolytic reaction products were further analyzed using polyacrylamide gel electrophoresis (15%). Lane 1: 86-mer circular **Oligonucleotide 2** alone, Lane 2: 86-mer circular **Oligonucleotide 2** in the presence of T7 endonucleaese I, Lane 3: 86-mer circular **Oligonucleotide 3** alone, Lane 4: 86-mer circular **Oligonucleotide 3** in the presence of T7 endonucleaese, Lane 5: 86-mer circular **Oligonucleotide 1** alone, Lane 6: 86-mer circular **Oligonucleotide 3** in the presence of T7 endonucleaese I.

5.2.6 Position Dependence of C3-Spacer Modification on the Inhibitory Efficiency of Topoisoemrase I

With the purpose of determining the correlation between position of C3-spacers in the dumbbell-shaped structures and inhibitory efficiency of human Topo I, additional oligonucleotides containing C3-spacers at certain different locations were subsequently synthesized (**Table 5-2**) and examined. As shown in **Table 5-1**, when a C3-spacer appeared near the Topo I-cutting site, the inhibitory efficiencies of the corresponding oligonucleotides decreased progressively (IC_{50} of **Oligonucleotide 4** = 89 nM; IC_{50} of **Oligonucleotide 5** = 52 nM while IC_{50} of **Oligonucleotide 1** = 37 nM).

This happened most likely because a C3-spacer, when it occurred near the cutting site of Topo I, could disturb the cutting-action of this DNA relaxing enzyme. On the other hand, when a C3-spacer emerged far off from Topo I-cutting site, the corresponding oligonucleotides displayed low inhibitory efficiency as well (IC_{50} of **Oligonucleotide 6** = 63 nM; IC_{50} of **Oligonucleotide 7** = 93 nM) as compared to **Oligonucleotide 1**. These decreases of inhibitory efficiency could be caused by that when these C3-spacers are located far away from the Topo I-cutting site, they will be in a position close to the second binding site of Topo I (**Figure 5-5**), which could consequently prevent the binding of this nuclear enzyme to the corresponding oligonucleotides.

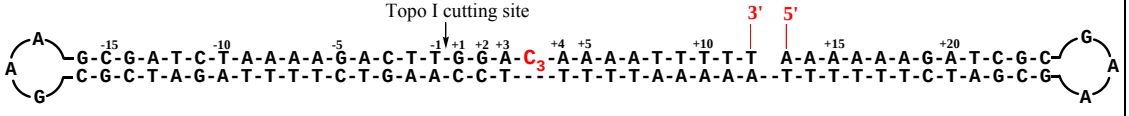
Table 5-1. Inhibitory efficiency (IC_{50}) of some C3-spacer-containing oligonucleotides on the activity of human Topo I.[a]

Names of oligonucleotides	IC_{50} (nM)
Oligonucleotide 4 (C3-spacer between G_{+1} to G_{+2})	89
Oligonucleotide 5 (C3-spacer between G_{+2} to A_{+3})	52
Oligonucleotide 1 (C3-spacer between A_{+3} to A_{+4})	37
Oligonucleotide 6 (C3-spacer between A_{+4} to A_{+5})	63
Oligonucleotide 7 (C3-spacer between A_{+5} to A_{+6})	93

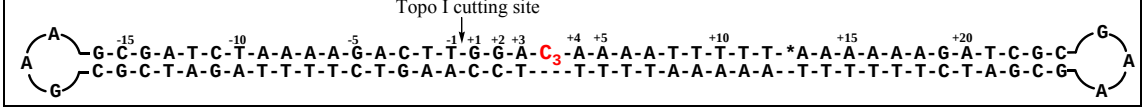
[a] See **Table.5-2** for detailed sequences of these oligonucleotides.

Table 5-2. Sequences and C3-spacer modifications of oligonucleotides prepared during this study:[b]

Oligonucleotide L (Linear Precursor):

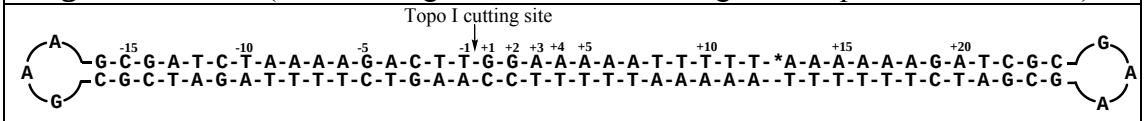


Oligonucleotide 1 (C3-spacer-containing circular oligonucleotide):

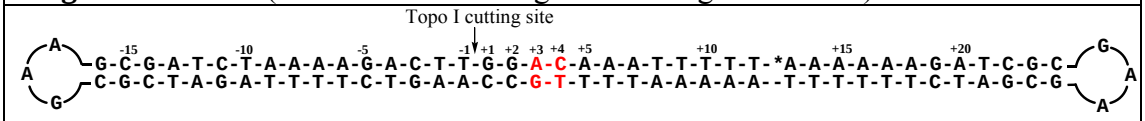


**** C3-spacer is positioned between A_{+3} and A_{+4}***

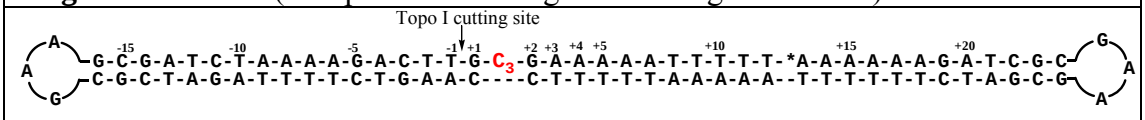
Oligonucleotide 2 (Circular oligonucleotide containing no C3-spacer modification):



Oligonucleotide 3 (Mismatch-containing circular oligonucleotide):

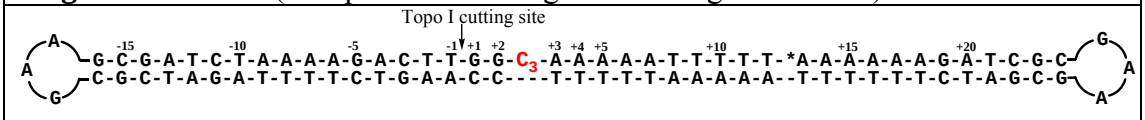


Oligonucleotide 4 (C3-spacer-containing circular oligonucleotide):



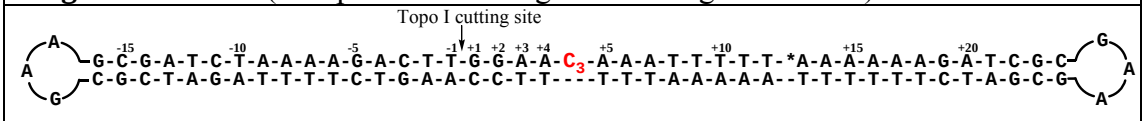
**** C3-spacer is positioned between A_{+1} and A_{+2}***

Oligonucleotide 5 (C3-spacer-containing circular oligonucleotide):



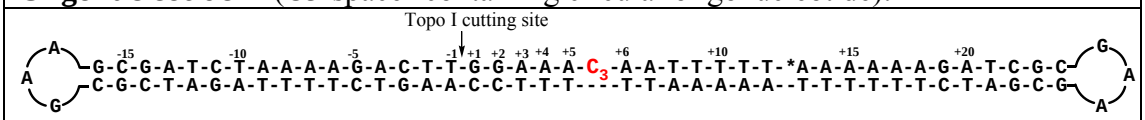
**** C3-spacer is positioned between A_{+2} and A_{+3}***

Oligonucleotide 6 (C3-spacer-containing circular oligonucleotide):



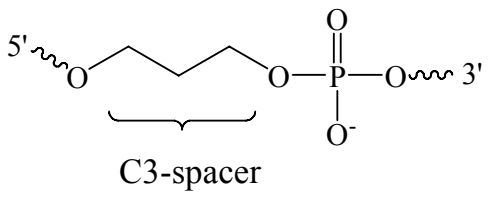
**** C3-spacer is positioned between A_{+4} and A_{+5}***

Oligonucleotide 7 (C3-spacer-containing circular oligonucleotide):



**** C3-spacer is positioned between A_{+5} and A_{+6}***

[b]



5.3 Gap-Containing Unimolecular Oligonucleotides as Topoisomerase I Inhibitors

The innate substrates of DNA topoisomerase I *in vivo* are positive and negative supercoiled DNA formed from the processes of replication, transcription, and chromatin condensation. It has been demonstrated in the past years, in addition to these supercoiled structures, that a few linear duplex oligonucleotides could act as substrates of eukaryotic topoisomerase I *in vitro*.[185-189] One of such duplex sequences (**Duplex 1**) that contains topoisomerase I-binding and cutting sites are illustrated in **Figure 5-14**.[192-197] Subsequent investigation in this research field demonstrated that when a gap was present in the topoisomerase I-binding sequence near its cutting site, one strand of the duplex oligonucleotide substrates could generate covalent linkages with the enzyme in an irretrievable manner.[204-205] These discoveries have given us the inspiration to look recently at the possibility of using gap-containing unimolecular oligonucleotides as irreversible inhibitors of human topoisomerase I. Here we report the results of our examination on the inhibitory effect of certain oligonucleotides on the relaxation reaction of negatively supercoiled DNA catalyzed by human topoisomerase I.

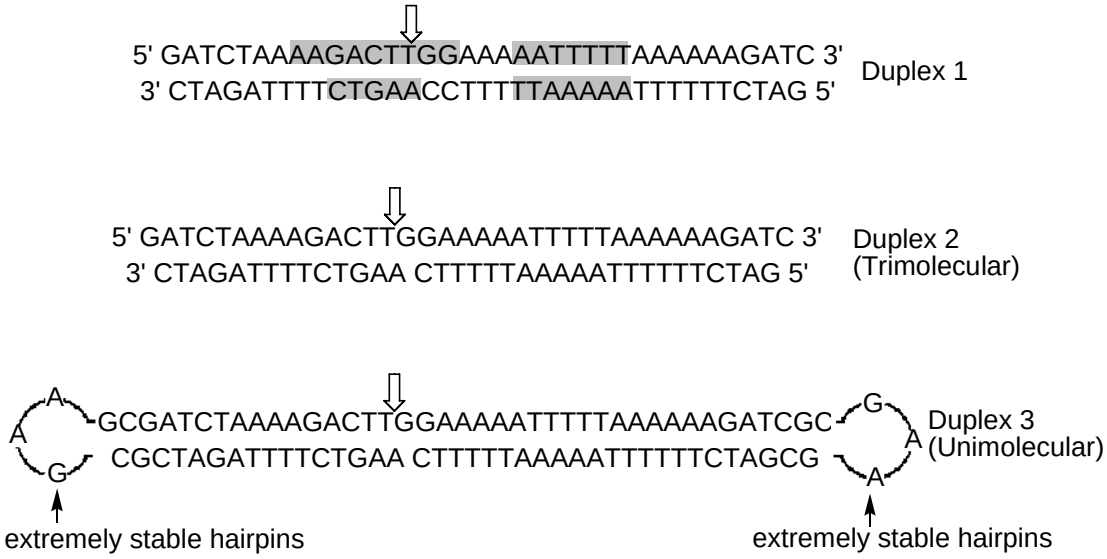


Figure 5-14. Sequences of oligonucleotides used in the study of Topoisomerase I Inhibitors. The shadowed tracts in **Duplex 1** denote the topoisomerase I-binding tracts, and the cutting site by topoisomerase I is indicated by ⇩.

5.3.1 Design Strategy of Gap-containing Oligonucleotides as Topoisomerase I Inhibitors

A gap-containing unimolecular oligonucleotide **Duplex 3** in **Figure 5-14** was also designed that possessed two "extraordinarily thermostable hairpins"[201] at the two end of the duplex structure according to our early investigation. These hairpin structures could in theory lead to an increase of the thermal stability of **Duplex 3**. In addition, it was anticipated that topoisomerase I could bind to **Duplex 3** and further cause a scission on one of the two strands at the designed cutting site if topoisomerase I and **Duplex 3** are allowed to interact with each other. After a nick is generated in **Duplex 3** by topoisomerase I, a religation reaction between the cut fragments might not be able to take place because the presence of a gap near the cutting site could cause improper conformational alteration, thus leading to chemically irreversible damage to the enzyme[204, 205] (**Figure 5-15**).

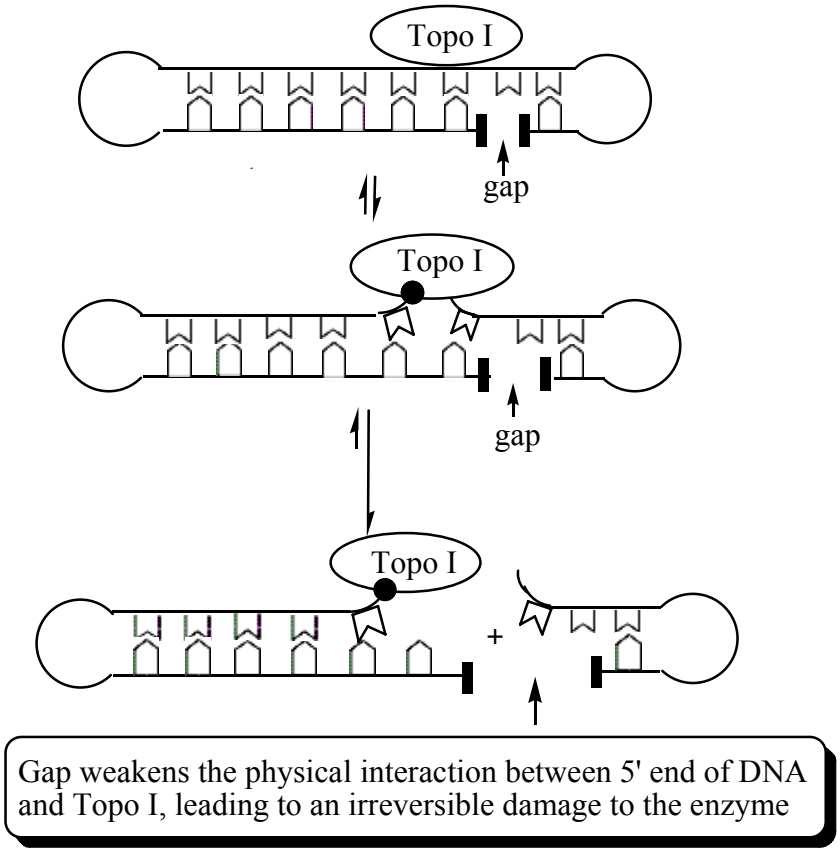


Figure 5-15. Illustration of possible mechanism for gap-containing oligonucleotides as Topoisomerase I Inhibitors.

5.3.2 Examination of Inhibitory Effect of Gap-containing Oligonucleotides as Topoisomerase I Inhibitors

Two different types of tests about the inhibitory effects of **Duplex 3** on human topoisomerase I were performed during our investigations. **Duplex 3** was preincubated with topoisomerase I for 3 minutes prior to the addition of plasmid DNA in the first types of tests (preincubaton test) while in the second type of tests, topoisomerase I, pBR322 and **Duplex 3** were mixed together simultaneously (tests without preincubation). During the preincubation tests, **Duplex 3** was expected to have enough time to bind and further generate irreversible linkages with topoisomerase I without the inference of pBR322. As shown in **Figure 5-16**, relaxation of negatively supercoiled pBR322 by topoisomerase I was near completion

in the absence of **Duplex 3** (Lane 2). With the addition and increase of concentration of **Duplex 3**, on the other hand, the inhibitory effect of this dumbbell-shaped structure on the relaxation reaction of supercoiled pBR322 increases (Lane 3 to Lane 8). The IC_{50} value of **Duplex 3** under our reaction conditions was 72.3 nM (**Figure 5-17**).

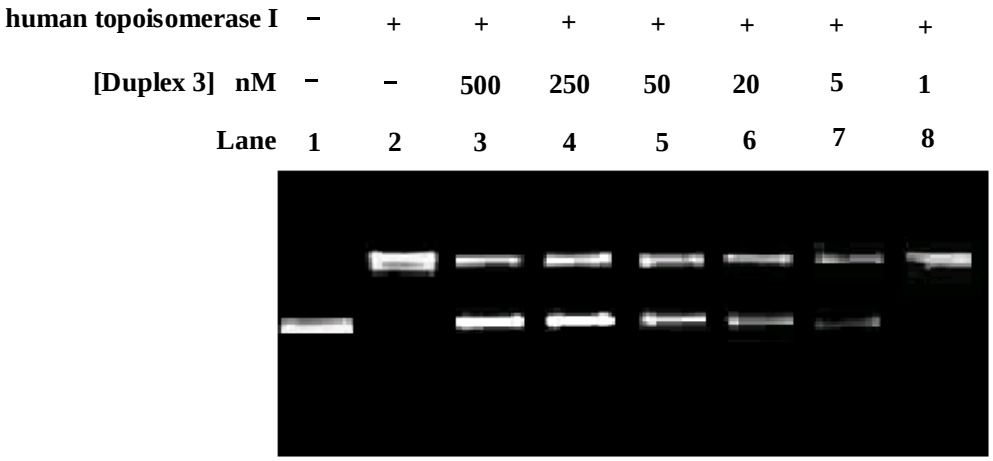


Figure 5-16. Agarose gel electrophoretic analysis of inhibitory effect of **Duplex 3** on human topoisomerase I. **Duplex 3** and 1 unit human topoisomerase I were incubated in a buffer of 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM PMSF and 1.0 mM mercaptoethanol in a total volume of 20 μL at 25 °C for 3 min. When this preincubation was over, pBR322 was added and the new mixtures were adjusted to contain 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 1 mM PMSF, 1.0 nM mercaptoethanol, 400 ng pBR322, 1 unit of human topoisomerase I and different concentrations of **Duplex 3**, which was further allowed to react at 37 °C for 20 mins. Concentrations of **Duplex 3** in the samples from Lane 3 to Lane 8 were 500 nM, 250 nM, 50 nM, 20 nM, 5 nM and 1 nM respectively.

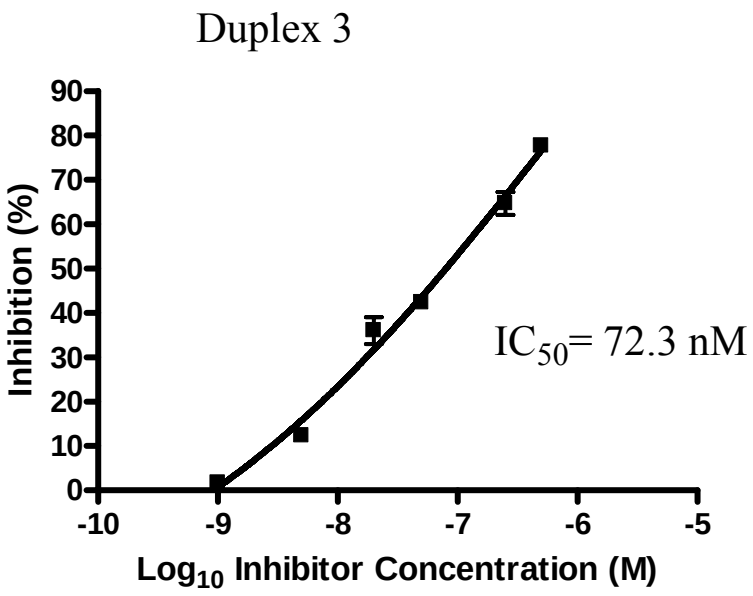


Figure 5-17. Correlations between percent inhibition on topoisomerase I activity and concentration of **Duplex 3**. Percentage of relaxation was defined as the ratio of band density of relaxed DNA over those of relaxed DNA plus supercoiled DNA [relaxed DNA/ (relaxed DNA + supercoil DNA)] while (100%- percentage of relaxation) was taken as the percent inhibition of topoisomerase I activity by oligonucleotides. The DNA bands were analyzed using Gel Documentation System (G:Box HR, Syngnene, Cambridge, UK) equipped with Gene Tools Software. IC_{50} value is calculated by GraphPad Prism 4 software.

As a comparison test, the inhibitory effect of **Duplex 2** was also examined in our studies, which contains the same duplex sequence as **Duplex 3** does (**Figure 5-14**). Different from **Duplex 3**, on the other hand, **Duplex 2** is a trimolecular complex rather than a unimolecular entity and does not possess two "extremely stable hairpins" in its structure. As shown in **Figure 5-18**, this trimolecular structure exhibited inhibitory effect on the relaxation reaction of pBR322 by topoisomerase I as well. The IC_{50} value of **Duplex 2** is 568 nM, which is about 8 times greater than the one obtained when **Duplex 3** was used (72.3 nM). The difference of IC_{50} values between **Duplex 2** and **Duplex 3** could be attributed to that the unimolecular structure of **Duplex 3** has a higher melting point and generates duplex structure more efficiently than the trimolecular structure of **Duplex 2** does.

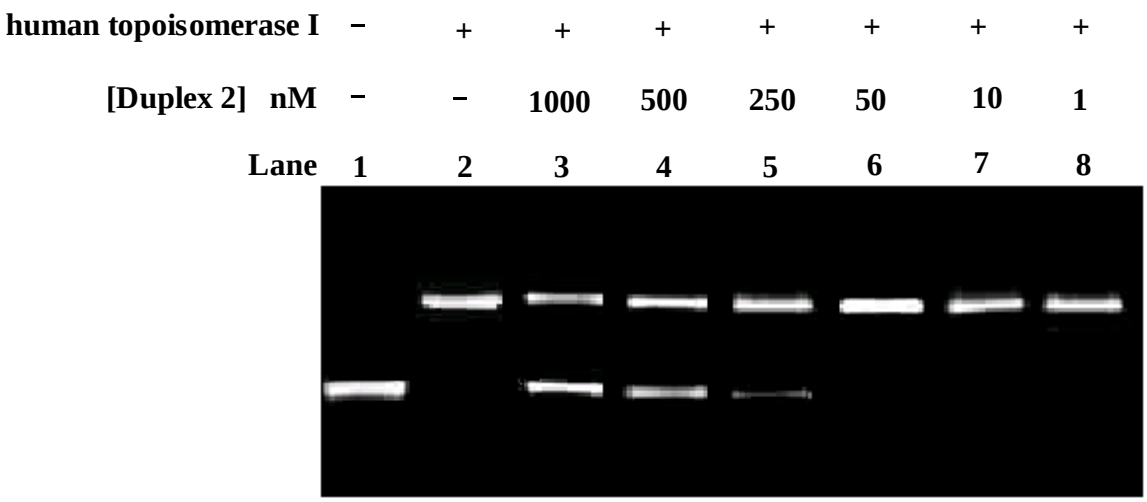


Figure 5-18. Agarose gel electrophoretic analysis of inhibitory effect of **Duplex 2** on human topoisomerase I. The same procedures as those prepared for the samples in **Figure 5-16** were used except for replacing **Duplex 3** with **Duplex 2**. Concentrations of **Duplex 2** in the samples from Lane 3 to Lane 8 were 1000 nM, 500 nM, 250 nM, 50 nM, 10 nM, and 1 nM respectively.

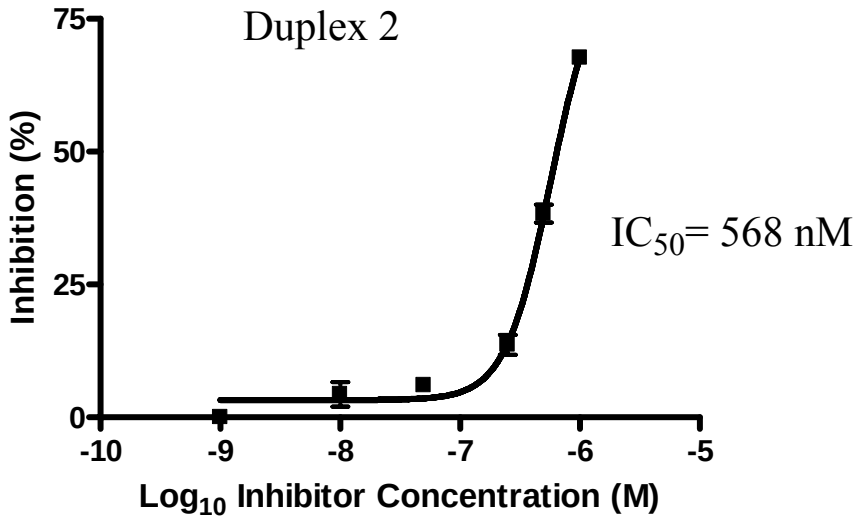


Figure 5-19. Correlations between percent inhibition on topoisomerase I activity and concentration of **Duplex 2**. IC_{50} value is calculated by GraphPad Prism 4 software.

Duplex 1 is a 36 mer linear duplex sequence containing topoisomerase I-binding and cutting sites (**Figure 5-14**). Different from **Duplex 2** and **Duplex 3**, this duplex structure contains perfectly matched base pairs in its structure. As control experiments, the inhibitory effect of **Duplex 1** on the activity of topoisomerase I was

examined during our investigation. As shown in **Figure 5-20**, this bimolecular complex displayed much lesser inhibitory efficiency than **Duplex 2** and **Duplex 3** do and showed an IC_{50} value of about 1.4 μM. The lower inhibitory efficiency of **Duplex 1** could possibly be due to that the perfectly matched sequence of **Duplex 1** is not capable of forming an irreversible linkage with topoisomerase I during the strand scission and rejoining process catalyzed by topoisomerase I .

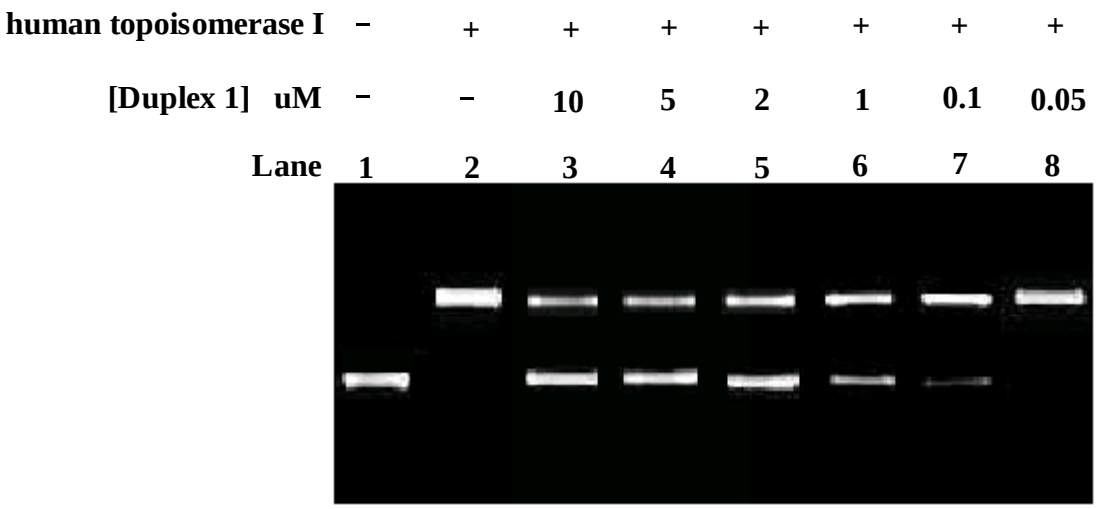


Figure 5-20. Agarose gel electrophoretic analysis of inhibitory effect of **Duplex 1** on human topoisomerase I. The same procedures as those prepared for the samples in **Figure 5-16** were used except for replacing **Duplex 3** with **Duplex 1**. Concentrations of **Duplex 1** in the samples from Lane 3 to Lane 7 were 10,000 nM, 5,000 nM, 2,000 nM, 1,000 nM, 100 nM and 50 nM respectively.

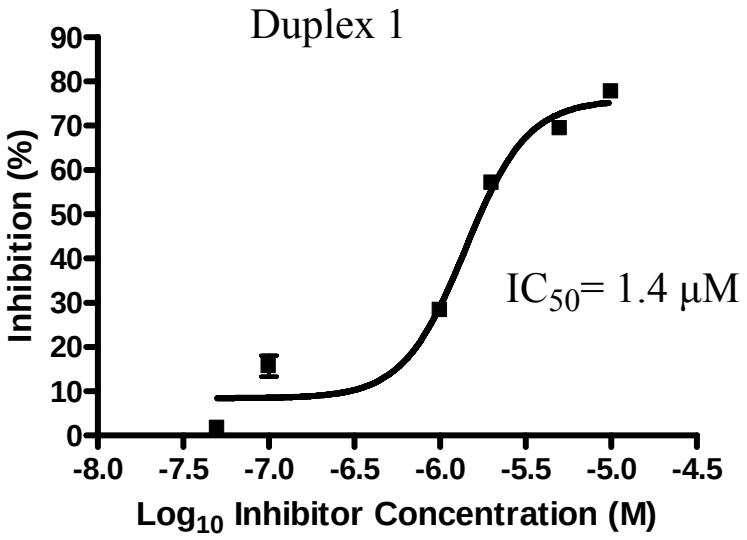


Figure 5-21. Correlations between percent inhibition on topoisomerase I activity and concentration of **Duplex 1**. IC_{50} value is calculated by GraphPad Prism 4 software.

In order to make comparison with the preincubation studies (**Figure 5-16**), new tests without preincubation were further carried out during our investigations in which **Duplex 3**, topoisomerase I and pBR322 were mixed together simultaneously (**Figure 5-22**). Under this condition, **Duplex 3** displayed an IC_{50} value of 1.04 μM (**Figure 5-23**), which is relatively high as compared with the one obtained when **Duplex 2** was preincubated with topoisomerase I for three minutes (IC_{50}= 72.3 nM). Competitions between **Duplex 3** and pBR322 in their binding to topoisomerase I could be the cause of low efficiency of **Duplex 3** in these inhibitory tests.

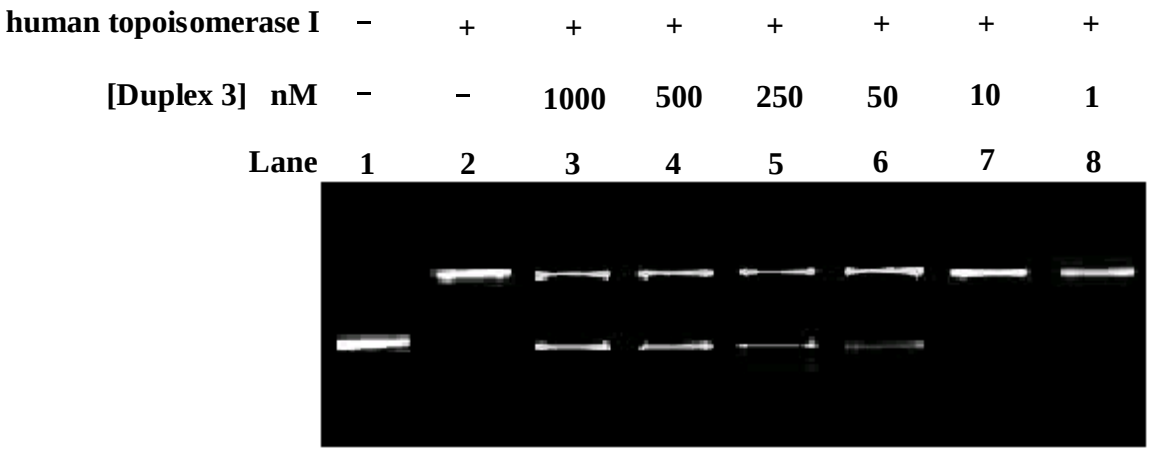


Figure 5-22. Agarose gel electrophoretic analysis of inhibitory effect of **Duplex 3** on human topoisomerase I without preincubation. A mixture containing 10 mM TRIS (pH 7.5), 100 mM NaCl, 400 ng pBR322, 1 units of human topoisomerase I and different concentrations of **Duplex 3** were prepared and further incubated at 37 °C for 20 min. Concentration of **Duplex 3** in the samples from Lane 3 to Lane 8 were 1000 nM, 500 nM, 250 nM, 50 nM, 10 nM and 1 nM respectively.

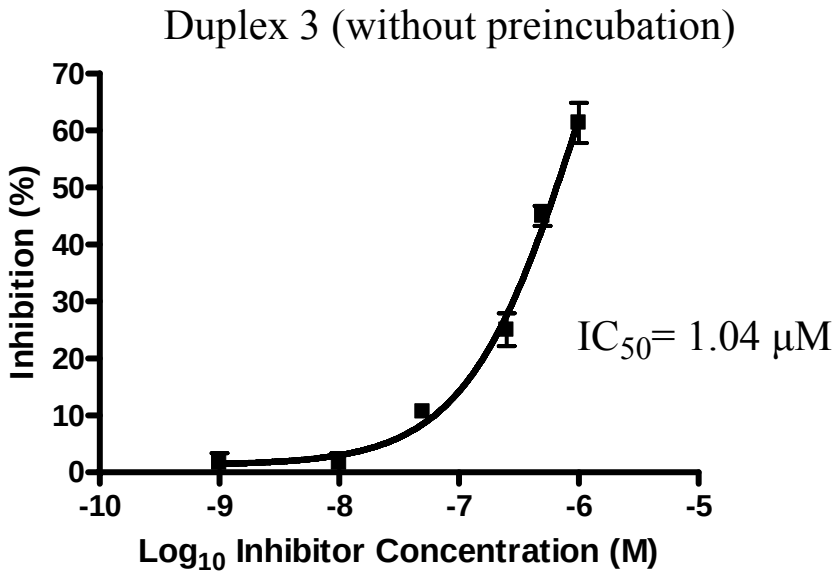


Figure 5-23. Correlations between percent inhibition on topoisomerase I activity and concentration of **Duplex 3** without preincubation. IC_{50} value is calculated by GraphPad Prism 4 software.

5.4 Conclusions

It was demonstrated in our studies that some C3-spacer-containing dumbbell-shaped oligonucleotides not only exhibited high inhibitory efficiencies on the activity of human topoisomerase I (e.g. IC_{50} = 36.6 nM for **Oligonucleotide 1**), but also found resistant to the hydrolysis by T7 endonuclease I. In addition, gapped unimolecular oligonucleotide containing topoisomerase I-binding site displayed an IC_{50} value of 72.3 nM, which is efficient than most of the organic inhibitors (e.g. IC_{50} of camptothecin = 700 nM) in its action on topoisomerase I.

Recent development demonstrated that oligonucleotides can be introducd into eukaryotic cells with reduced cytotoxicity, we expect the newly designed dumbbell-shaped circular oligonucleotides containing internal C3-spacers as well as gapped unimolecular oligonucleotide could act as potential DNA-based therapeutical agents that target topoisomerase I. It is also our anticipation that the outcomes of these studies could have certain implications in the future design of new inhibitors of DNA topoisomerase I.

Chapter 6

Materials and Methods

6.1 Materials

6.1.1 Oligonucleotides

All oligonucleotides used in our studies are purchased from Sigma-Proligo (Singapore) which provided custom oligonucleotides synthesis.

Oligonucleotides are usually supplied as a lyophilized powder that could be re-suspended in water and diluted to suitable concentration for reactions. All oligonucleotides samples should be stored at - 20°C.

Common synthetic oligonucleotides are provided with 5' end hydroxyl group which is suitable for 5' end labeling reactions. Oligonucleotides used for concentration controlled circulization reactions in our studies are purchased with 5' end phosphorylated. Oligonucleotides used for DNA topoisomerase I inhibitors (Chapter 5) are purchased with C3-spacer modification.

6.1.2 Enzymes

Enzymes used in our studies are all from commercial sources.

a) T4 polynucleotide kinase

T4 polynucleotide kinase used in our studies is purchased from New England Biolabs (Cat. No. M0201L, 2500 units per package, 10 units/μL in 10 mM Tris-HCl, 50 mM KCl, 0.1 μM ATP, 1 mM Dithiothreitol, 0.1 mM EDTA and 50% Glycerol at pH 7.4 @ 25°C). The enzyme should be stored at - 20°C. It is purified from an *E. coli*

strain carrying a cloned of T4 polynucleotide kinase gene.

T4 polynucleotide kinase catalyzes the transfer or exchange of phosphate from the γ position of ATP to the 5′ -hydroxyl terminus of polynucleotides (double-and single-stranded DNA and RNA) and nucleoside 3′-monophosphates. It also catalyzes the removal of 3′-phosphoryl groups from 3′-phosphoryl polynucleotides, deoxynucleoside 3′-monophosphates and deoxynucleoside 3′-diphosphates.

One unit of T4 polynucleotide kinase incorporates 1 nmol of acid-insoluble [^{32}P] in a total reaction volume of 50 µl in 30 minutes at 37°C in 1X T4 Polynucleotide Kinase Reaction Buffer with 66 µM [γ-^{32}P] ATP (5 x 10^6 cpm/µmol) and 0.26 mM 5′-hydroxyl-terminated salmon sperm DNA. The enzyme was supplied with 10 X reaction buffer (0.7 M Tris-HCl, pH 7.6, 100 mM $MgCl_2$, 50 mM Dithiothreitol).[209]

b) T4 DNA ligase

T4 DNA ligase used in our studies is purchased from New England Biolabs (Cat. Number: M0202L, 100,000 units per package, 400 units/µL in 10 mM Tris-HCl, 50 mM KCl, 1 mM DTT, 0.1 mM EDTA and 50% Glycerol, pH 7.4 @ 25°C). The enzyme should be stored at -20 °C.

T4 DNA ligases catalyze the formation of a phosphodiester bond between juxtaposed 5' phosphate and 3' hydroxyl termini in duplex DNA or RNA. This enzyme will join blunt end and cohesive end termini as well as repair single stranded nicks in duplex DNA, RNA or DNA/RNA hybrids.

One unit of T4 DNA ligase gives 50% ligation of HindIII fragments of λ DNA (5′ DNA termini concentration of 0.12 µM, 300 µg/ml) in a total reaction volume of 20 µl in 30 minutes at 16°C in 1X T4 DNA Ligase Reaction Buffer. The enzyme can

also catalyze the ligation at room temperature. For cohesive ends, use 1 µl of T4 DNA Ligase in a 20 µl reaction for 10 minutes. For blunt ends, use 1 µl of T4 DNA Ligase in a 20 µl reaction for 2 hours or 1 µl high concentration T4 DNA Ligase for 10 minutes. The enzyme was supplied with 10 X reaction buffer (0.5 M Tris-HCl, 100 mM $MgCl_2$, 10 mM ATP, 100 mM Dithiothreitol, pH 7.5 @ 25°C).[210]

c) Exonuclease I

Exonuclease I used in our studies is purchased from New England Biolabs (Cat. Number: M0293L, 15,000 units per package, 20 units/µL in 10 mM Tris-HCl, 100 mM NaCl, 5 mM 2-Mercaptoethanol, 0.5 mM EDTA, 100 µg/ml BSA and 50% Glycerol, pH 7.5 @ 25°C). The enzyme should be stored at -20°C.

Exonuclease I was purified from an *E. coli* strain that carries the cloned Exo I gene from *E. coli* NM554. It catalyzes the removal of nucleotides from single-stranded DNA in the 3' to 5' direction and degrades excess single-stranded primer oligonucleotides from a reaction mixture containing double-stranded extension products.

One unit is defined as the amount of enzyme that will catalyze the release of 10 nmol of acid-soluble nucleotide in a total reaction volume of 50 µl in 30 minutes at 37°C in 1X Exonuclease I Reaction Buffer with 0.17 mg/ml single-stranded [^{3}H]-DNA. The enzyme was supplied with 10 X reaction buffer (670 mM Glycine-KOH, 67 mM $MgCl_2$, and 100 mM 2-Mercaptoethanol, pH 9.5 @ 25°C).[211]

d) Exonuclease VII

Exonuclease VII used in our studies is purchased from GE-Healthcare (Cat. Number: 70082Y, 200 units per package, 10 units/µL in 50 mM Tris-HCl, pH 8.0,

200 mM NaCl, 0.1 mM EDTA, 10 mM DTT and 50% glycerol.). The enzyme should be stored at -20°C.

Exonuclease VII is purified from *E. coli* strain containing overproducing clones encoding both the large (xseA) and small (xseB) subunits of Exonuclease VII. A strict single-strand directed enzyme with 5' to 3' and 3' to 5'- exonuclease activities making it the only bi-directional *E. coli* exonuclease with single-strand specificity. One unit catalyzes the release of 1 nmol of acid-soluble nucleotide in 30 min at 37°C under standard conditions.[212]

e) DNase I

DNase I used in our studies is purchased from New England Biolabs (Cat. Number: M0303L, 5,000 units per package, 2 units/µL in 10 mM Tris-HCl, 2 mM $CaCl_2$, and 50% Glycerol, pH 7.6 @ 25°C). The enzyme should be stored at -20 °C.

DNase I is an endonuclease that nonspecifically cleaves DNA to release di-, tri- and oligonucleotides products with 5′ -phosphorylated and 3′ -hydroxylated ends. DNase I acts on single- and double-stranded DNA, chromatin and RNA: DNA hybrids.

One unit is defined as the amount of enzyme which will completely degrade 1 µg of pBR322 DNA in 10 minutes at 37 °C in DNase I Reaction Buffer. Complete degradation is defined as the reduction of the majority of DNA fragments to tetranucleotides or smaller. The enzyme was supplied with 10 X reaction buffer (100 mM Tris-HCl, 25 mM $MgCl_2$ and 5 mM $CaCl_2$).[213]

f) T7 endonuclease I (repair protein)

T7 Endonuclease I used in our studies is purchased from New England Biolabs (Cat. Number: M0302L, 1,250 units per package, 10 units/µL in 20 mM Tris-HCl, 200 mM NaCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 50% Glycerol and 0.15% Triton X-100, pH 7.5 @ 25°C). The enzyme should be stored at -20 °C.

T7 Endonuclease I recognizes and cleaves non-perfectly matched DNA, cruciform DNA structures, Holliday structures or junctions, heteroduplex DNA and more slowly, nicked double-stranded DNA. The cleavage site is at first, second or third phosphodiester bond that is 5′ to the mismatch. The protein is the product of T7 gene 3.

One unit is defined as the amount of enzyme required to convert > 90% of 1 µg of supercoiled cruciform pUC(AT) to > 90% linear form in a total reaction volume of 50 µl in 1 hour at 37°C. The enzyme was supplied with 10 X reaction buffer (100 mM Tris-HCl, 500 mM NaCl, 100 mM $MgCl_2$, and 1 0mM Dithiothreitol, pH 7.9 @ 25°C).[214]

g) T7 Exonuclease

T7 Exonuclease used in our studies is purchased from New England Biolabs (Cat. Number: M0263L, 5,000 units per package, 10 units/µL in 10 mM Tris-HCl, 5 mM Dithiothreitol, 0.1 mM EDTA and 50% Glycerol, pH 8.0 @ 25°C). The enzyme should be stored at -20 °C.

T7 Exonuclease is purified from an *E. coli* strain containing a TYB12 intein fusion. It acts in the 5' to 3' direction, catalyzing the removal of 5' mononucleotides from duplex DNA. T7 Exonuclease is able to initiate removal of nucleotides from the 5' termini or at gaps and nicks of double-stranded DNA.[216] It will degrade both 5'

phosphorylated or 5' dephosphorylated DNA. It has also been reported to degrade RNA and DNA from RNA/DNA hybrids in the 5' to 3' direction but is unable to degrade either double-stranded or single-stranded RNA. The protein is the product of T7 gene 6.[217]

One unit is defined as the amount of enzyme required to produce 1 nmol of acid-soluble deoxyribonucleotide from double-stranded DNA in a total reaction volume of 50 µl in 30 minutes at 25°C in 1X Reaction buffer with 0.15 mM sonicated duplex [^{3}H]-DNA. The enzyme was supplied with 10 X reaction buffer (200 mM Tris-acetate, 500 mM potassium acetate, 100 mM Magnesium Acetate and 10 mM Dithiothreitol).[215]

h) Mse I (restriction endonuclease)

Mse I is a unique restriction endonuclease purifed from *E. coli* strain that carries the MseI gene from *Micrococcus* species. It recognizes the double stranded sequence 5' TTAA 3' and cleaves between the two T residues.[219]

Mse I used in our studies is purchased from New England Biolabs (Cat. Number: R0525L, 2,500 units per package, 10 units/µL in 10 mM Tris-HCl 50 mM NaCl, 1 mM Dithiothreitol, 0.1 mM EDTA, 200 µg/ml BSA and 50% Glycerol, pH 7.4 @ 25°C). The enzyme should be stored at -20 °C.

One unit is defined as the amount of enzyme required to digest 1 µg of λ DNA in 1 hour at 37°C in a total reaction volume of 50 µl. The enzyme was supplied with 10 X reaction buffer (100 mM Tris-HCl, 500 mM NaCl, 100 mM $MgCl_2$, and 10 mM Dithiothreitol).[218]

i) Recombinant Human Topoisomerase I (wild type protein)

Recombinant Human Topoisomerase I used in our studies is purchased from Topogen (Catalog No. 2005H-RC1, 500 units per package, 8 units /μL in 20mM NaH_2PO_4 at pH 7.4, 300mM NaCl and 500mM Imidazole). The enzyme could be stored at 4 °C.

Topoisomerase I has been overexpressed in baculovirus and affinity purified as a single band on SDS-PAGE of 100kDa. The final preparation of enzyme is highly active in relaxation of plasmid DNA. One unit of Topoisomerase I can relax 0.25 μg (250 ng) of supercoiled plasmid DNA in 30 min at 37 °C. Relaxation assays were carried out in a final volume of 25 μl in Topisomerase I reaction buffer (10X reaction buffer): 100 mM Tris-Cl at pH 7.9, 1.5 M NaCl, 1 % BSA, 1 mM Spermidine, 50 % glycerol.[220]

6.1.3 pBR322 DNA

pBR322 used in our studies is purchased from New England Biolabs (Cat. Number: N3033L, 1,000 μg /ml in 10 mM Tris-HCl, 1 mM EDTA, pH 8.0 @ 25°C). This plasmid DNA vendor should be stored at -20°C.

pBR322 DNA is isolated from *E. coli* ER2272 by a standard plasmid purification procedure. It is a commonly used plasmid cloning vector in *E. coli*. [222] The molecule is a double-stranded circle 4,361 base pairs in length. [223] pBR322 contains the genes for resistance to ampicillin and tetracycline, and can be amplified with chloramphenicol. The molecular weight is 2.83 x 10^6 daltons.[221]

6.1.4 Buffer

a) TBE buffer

TBE buffer used in our studies are purchased from 1st Base Pte Ltd. (Cat.Number: 3010-10 X 1L, 1L per package). It is supplied in 10 times concentration solution which contains 1M Tris, Borate Acid and EDTA (pH 8.3).

b) TAE buffer

TAE buffer used in our studies are purchased from 1st Base Pte Ltd. (Cat.Number: 3000-10 X 1L, 1L per package). It is supplied in 10 times concentration solution which contains 1M Tris, Acetate Acid and EDTA (pH 8.0).

c) Tris-HCl

Tris-HCl buffer used in our studies are purchased from 1st Base Pte Ltd. (Cat.Number: 1415 -1L, 1L per package, contains 1M Tris (pH 7.0).

d) MES

MES used in our studies are purchased from Sigma-Aldrich (Cat. Number: M3671-50G) in solid state. It was dissolved in deionized water at desired concentration and adjusted to appropriate pH value by using 1M NaOH or 1M HCl.

e) HEPES

HEPES used in our studies are purchased from Sigma-Aldrich (Cat. Number: H3375-25G) in solid state. It was dissolved in deionized water at desired concentration and adjusted to appropriate pH value by using 1M NaOH or 1M HCl.

6.2 Methodology

6.2.1 5’ End labeling of DNA (T4 Polynucleotide Kinase method)

End labeling is a rapid and sensitive method in visualizing small amount of DNA. it includes 3’ End labeling and 5’ end labeling which introducing labeling groups onto 3’ or 5’ end of target sequences via enzyme catalytic reaction respectively. Compare to the rapidly developed nonradioactive method, radioactive labeling is still our first choince because 1) High sensitivity, the radioactive method enable only several molecules be detected, it always thousands time or more sensitive than nonradioactive methods; 2) the radioactive atom used in end labeling possessing same chemical and physical properties as their non-radioactive isotopes. ^{32}P 5’ end labeling provides us a possibility to study the oligpdeoxyribonucleotide in its native properties, especially in the ligation reaction of the 3’ and 5’ end together forming a normal phosphate bond.[224]

Polynucleotide kinase was commonly used in 5’ end labeling of single strand oligonucleotides. The standard reaction was conducted as following described:

5’ hycroxyl oligonucleotide (100 μM)	2 μL
10 X reaction buffer	2 μL
T4 PNK (10 units/ μL)	1 μL
[γ-^{32}P] ddATP (10 μCi/ μL)	2 μL
H_2O	13 2 μL

10 X reaction buffer including: 0.7 M Tris-HCl, pH 7.6, 100 mM MgCl, 50 mM Dithiothreitol.

The reaction mixture was incubated at 37 °C for 1 hour. 10 μL denaturing loading buffer was added before loading on a 20% (19:1) denature PAGE for further purification.

6.2.2 Polyacrylamide Gel Electrophoresis (PAGE)

Two types of polyacrylamide gels are in common use: One is nondenaturing polyacrylamide gel for the separation and purification of fragments of double-stranded DNA or other type secondary structure of DNA. The other is denaturing polyacrylamide gel for the separation and purification of single-stranded fragments of oligonucleotides. Preparations of denaturing and nondenaturing gels are similar.

Acrylamide stock solution (20% (19:1), 7M Urea). [224]

Acrylamide	190 g
N,N'-methylenebisacrylamide	10 g
10 X TBE buffer	100 mL
Urea	420 g
H_2O	To 1 L

10% Ammonium oersulfate solution was prepared via dissolving 1 g Ammonisum persulfate in 10 mL H2O. The solution may be stored at 4 °C for several weeks.

Place the required quantity of acrylamide solution in a clean beaker. Add 30 μL of TEMED (N,N,N',N'-tetramethylethylenediamine) to acrylamide solution, then 500 μL 10% (w/v) ammonium persulfate water solution. Mix the solution by stirring. Pour into the gel plate. The gel will be solidified in 15 mins.

6.2.3 DNA Purification (Desalting)

Oligonucleotides eluted from gel or produced in reactions are solved in buffers. Purifications of oligonucleotides are needed in most cases. Many methodologies have been developed to fulfill this purpose, such as dialysis, ion-exchange and solid phase extraction.

a) Solid phase extraction.

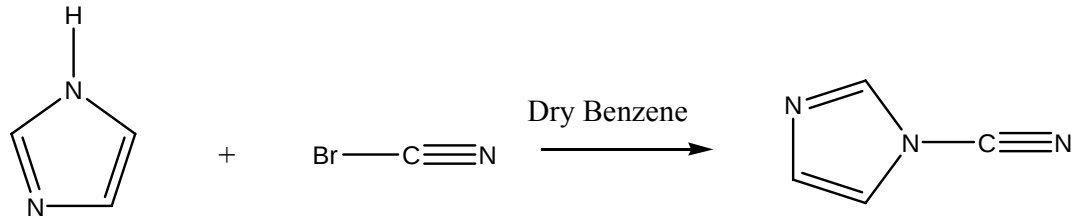
The common used is Waters Sep-Pak 1 mL (100 mg) C18 Cartridge (WAT023590). All solvents were applied to cartridge in a rush through method. New cartridge was activated by 500 μL CH_3CN for two times, washed by 500 μL H_2O for 3 times and loaded with 500 μL 1M Tris-HCl pH 7.0 buffer. The oligonucleotides solution was subsequently pushed through slowly at about 500 μL per minute. Cartridge containing sample was then washed with 500 μL water for 3-5 times. Purified oligonucleotides were eluted from column by 500 μL CH_3CN/H_2O (1:1 v/v). Fraction containing oligonucleotides was collected drop by drop and dried in SpinVac.

b) Ion-exchange purification

To purify the radio-labeled probes, the common used column is NAP-25 columns from GE Healthcare. The columns are pre-packed with Sephadex G-25 DNA Grade in distilled water (which contains 0.15% Kathon CG/ICP Biocide, used as a preservative) and can be stored at room temperature. They come in different sizes for purifying different volumes of samples: 0.5 ml (NAP-5), 1 ml (NAP-10), and 2.5 ml (NAP-25).

NAP columns are designed for the rapid and efficient desalting, buffer exchange and purification of DNA and oligonucleotides (equal or greater than 10 mers) utilizing gravity flow. The gel bed dimensions are 1.5 x 4.9 cm, the maximum sample volume is 2.5 ml, and the volume of eluted sample is 3.5 ml. The columns must be equilibrated with 25 ml 10 mM sodium phosphate buffer (pH 6.8), Once equilibrated, the probe is added in 2.5 ml and additional 3.5 ml of equilibration buffer is then added and 1.5 ml fractions are collected and dried in SpinVac.

6.2.4 Preparation of N-cyanoimidazole (Chapter 4)



A solution of 0.05 moles BrCN (Sigma-Aldrich, Cat. Number: 16774) in 25 ml dried benzene (Sigma-Aldrich, Cat. Number: 319953) was added dropwisely with stirring to a solution of 0.05 moles imidazole (Sigma-Aldrich, Cat. Number: I0125) in 50 mL dried benzene under 50 °C. After addition, the mixture was stirred at 50 °C for another 5 mins, and then cooled to 4 °C. The reaction was kept at 4 °C for 24 hours and the bright yellow solid was removed by filtration. A clear, colorless filtration was concentrated to dryness via rotary evaporator. The product, white crystalline solid was purified by sublimation. Yielding about 10% needle like colorless crystal and stored at -20°C in dark.[225]

6.2.5 Chemical ligation reactions of unimolecular G-quadruplex using N-cyanoimidazole (Chapter 4)

A stock solution of **Sequence 1** in 100 mM MES (pH 6.0) and 100 mM KCl were kept at 100 °C for 3 min followed by cooling the mixture to 25 °C over a period of 120 min to allow the formation of unimolecular G-quadruplexes. The reaction mixtures consisting of 1 μM of **Sequence 1** taken from the stock solution, 100 mM MES (pH 6.0), 100 mM KCl, 10 mM $MnCl_2$ and 10 mM N-cyanoimidazole in a total volume of 20 μL were then prepared and incubated at 25 °C for 6 h. The reaction was next terminated by the addition of loading buffer and the resultant ligation product was analyzed by polyacrylamide gel electrophoresis.

6.2.6 Self-cleavage reactions of Oligonucleotide 1 (Chapter 3)

All cleavable oligonucleotides were incubated at 20 °C in a buffer solution of 5mM HEPES (pH 7.0) and in the presence of 5 mM KCl for 2 h in order to allow the formation of proper G-quadruplex assemblies. The self-cleavage reactions of **Oligonucleotide 1** were initiated by adding $MgCl_2$ to the reaction mixture, which was then kept at 30 °C for an additional 2 h to allow the site-specific self-cleavage.

6.2.7 Fluorescence measurement (Chapter 2)

Fluorescence measurements were done on a Perkin-Elmer LS55 fluoremeter. All the emission spectra were taken by excitation at 480nm, which is the absorption maximum of Bodipy 493/503. All the measurements were done in 10 mM MES/HEPES buffer solutions with variation of pH value and data were recorded at 20 °C.

6.2.8 Thermal Stability Analysis of Oligonucleotides by UV spectroscopy (Chapter 2)

4 uM oligonucleotide were prepared in 10 mM Tris-HCl (pH 7.9) and 50 mM NaCl of total volume of 400 μL for its thermal stability analysis. The melting points of the oligonucleotides were determined by measuring the absorbance of the oligonucleotides at 260 nm by an UV spectrophotometer (*Nicolet Evolution 3000, Thermo Electron Corporation*) equipped with a thermoelectric cell holder. The rate of temperature increase was adjusted to 0.5 °C / min. The melting points of the oligonucleotides were determined by the inflection point on its melting point curves.

6.2.9 CD measurement (Chapter 3 & 4)

The samples were dissolved in 10 mM Tris-HCl buffer solution, various concentrations of metal ions and allowed to equilibrate overnight. CD spectra were recorded on a *Jasco J-810 spectropolarimeter* in a 1mm pathlengh cuvette. The wavelength scans were done at room temperature with a scan rate of 100nm/min. For thermal denaturation experiments, the temperature was changed from 20 °C to 90 °C at a rate of 2 °C/min and the CD at 295 nm was monitored.

6.2.10 Empirical estimation of duplex Melting temperature

Melting temperature (T_m) is the temperature at which an oligonucleotide duplex is 50% in single-stranded form and 50% in double-stranded form. The empirical estimation of T_M based on the nearest-neighbor two-state model, which is applicable to short DNA duplexes.

$$\text{Tm (°C)} = \frac{\triangle H°}{\triangle S° + R \ln[\text{oligo}]} - 273.15$$

where ΔH° (enthalpy) and ΔS° (entropy) are the melting parameters calculated from the sequence and the published nearest neighbor thermodynamic parameters, R is the ideal gas constant (1.987 $cal \cdot K^{-1} mole^{-1}$), [oligo] is the molar concentration of an oligonucleotide, and the constant of - 273.15 converts temperature from Kelvin to degrees of Celsius.

$$\triangle G = RT \ln\left(\frac{[\text{DNA}\bullet\text{primer}]}{[\text{DNA}]\bullet[\text{Primer}]}\right)$$

$$\triangle H = T\triangle S + RT \ln\left(\frac{[\text{DNA}\bullet\text{primer}]}{[\text{DNA}]\bullet[\text{Primer}]}\right)$$

$$\triangle H = \triangle G + T\triangle S$$

$$T = \frac{\triangle H - 3.4 \frac{kcal}{K\ mole}}{\triangle S + R \ln\left(\frac{1}{[primer]}\right)} + 16.6 \log_{10}([Na^+])$$

T_M depends on monovalent salt concentration ($[Na^+]$) of the solvent. The improved quadratic T_M salt correction function was reported by Owczarzy, R. et al. which is used in our studies.

$$\frac{1}{Tm(Na^+)} = \frac{1}{Tm(1M\ Na+)} + 4.29\ f(GC)-3.95)\ X\ 10^{-5}\ \ln[Na^+] + 9.40\ X\ 10^{-6}\ \ln^2\ [Na^+]$$

* where f(GC) is the fraction of GC base pairs.

6.2.11 General procure for Exonuclease VII hydrolysis (Chapter 4)

The identified circular product with a slower rate of mobility shift (Lane 9, **Figure 7**) was purified through polyacrylamide gel electrophoresis. A reaction mixture containing 1×exonuclease VII buffer (100 mM Tris-HCl, 100 mM potassium phosphate, 16.6 mM ethylenediamine tetraacetic acid and 20 mM 2-mercaptoethanol, pH 7.9), 20 units of exonuclease VII and the identified circular product or **Sequence 1** in a total volume of 20 μL was subsequently prepared and incubated at 37 °C for 2 h. This reaction product was next analyzed through polyacrylamide gel electrophoresis.

6.2.12 Partial hydrolysis of the identified circular product by DNAse I (Chapter 4)

A reaction mixture containing 1 x DNAse I buffer (100 mM sodium acetate and 5 mM magnesium sulfate, pH 5.0), 10 units DNAse I and the identified circular product in a total volume of 20 μL was incubated at room temperature for 30 min. This hydrolysis reaction was terminated by freezing the samples in liquid nitrogen and the reaction product was analyzed through polyacrylamide gel electrophoresis.

References:

1 Neidle, S. *Oxford Handbook of Nucleic Acid Structure*; Cambridge University Press, **1999**, pp. 39-74; Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag, New York, **1984**, pp. 9-13

2 Tobin, A. J. and Dusheck, J. *Asking About Life, 3rd Edition*, Thomson Brooks/Cole, **2005**, pp.54-56; 194-196.

3 Simonsson, T. *Biol. Chem.* **2001**, 382, 621-628.

4 Mills, M., Lacroix, L., Arimondo, P. B., Leroy, J.-L., Francois, J.-C.e, Klump, H. and Mergny, J.-L. *Curr. Med. Chem.* **2002**, 2, 627-644.

5 Belmont, P., Constant, J. F. and Demeunynck, M. *Chem. Soc. Rev.* **2001**, 30, 70-81.

6 Rhodes, D. and Giraldo, R. *Curr. Opin. Struct. Biol.* **1995**, 5, 311-322.

7 Lane, A. N. and Jenkins, T. C. *Curr. Org. Chem.* **2001**, 5, 845-869.

8 Patel, D. J., Bouaziz, S., Kettani, A. and Wang, Y. In Neidle, S. (Ed.), *Oxford Handb. Nucleic Acid Struct.* Oxford University Press, Oxford, UK, **1999**, pp. 389-453.

9 Shafer, R. H. *Prog. Nucleic. Acid. Re.* **1998**, 59, 55-94

10 Williamson, J. R. *Annu. Rev. Bioph. Biom.* **1994**, 23, 703-730

11 Wellinger, R. J. and Sen, D. *Eur. J. Cancer* **1997**, 33, 735-749.

12 Gilbert, D. E. and Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, 9, 305-314.

13 Mergny, J.-L.and Helene, C. *Nat. Med.* **1998**, 4, 1366-1367.

14 Arthanari, H. and Bolton, P. H. *Chem. Biol.* **2001**, 8, 221-230.

15 Pilch, D. S., Plum, G. E. and Breslauer, K. J. *Curr. Opin. Struct. Biol.* **1995**, 5, 334-342.

16 Sundquist, W. I. *Curr. Biol.* **1993**, 3, 893-895.

17 Rippe, K., Kuryavyi, V. V.,Westhof, E. and Jovin, T. M. Lilley, D. M. J., Heumann, H. and Suck, D. (Ed.) *Struct. Tools Anal. Protein-Nucleic Acid Complexes*, Birkhauser, Basel, Switz, **1992**, pp. 81-107.

18 Smith, F. W. and Feigon, J. *Nature* **1992**, 356, 164-168.

19 Jin, R., Breslauer, K. J., Jones, R. A. and Gaffney, B. L. *Science* **1990**, 250, 543-546.

20 Iulian, N., Rujan, J., Meleney, C. and Bolton, P. H. *Nucleic Acids Res.* **2005**, 33, 2022-2031.

21 Miyoshia, D., Nakaoa, A., Todaa, T. and Sugimotoa, N. *FEBS Letters* **2001**, 496, 128-133; Miyoshia, D., Nakaoa, A. and Sugimotoa, N., *Nucleic Acids Res.*, **2003**, 31, 1156-1163; Chen, F-M., *Biochemistry* **1992**, 31, 3769-3776; Smirnov, I. and Shafer, R. H., *J. Mol. Biol.* **2000**, 296, 1-5.

22 Bang, I. *Bioch. Ztschr*. **1910**, 26, 293

23 Gellert, M., Lipsett, M. N. and Davies, D. R. *Proc. Natl. Acad. Sci. USA* **1962**, 48, 2013-2018.

24 Pinnavaia, T. J., Marshall, C. L., Mettler, C. M., Fisk, C. L., Miles, H. T. and Becker, E. D. *J. Am. Chem. Soc.* **1978,** 100, 3625-3627.

25 Zhou, T., Chen, G., Wang, Y, Zhang, Q., Yang, M. and Li, T. *Nucleic Acids Res.* **2004**, 32, e173/1-e173/9.

26 Aboul-ela, F., Murchie, A. I. and Lilley, D. M. *Nature* **1992**, 360, 280-282; Aboul-ela, F., Murchie, A. I., Norman, D. G. and Lilley, D. M. *J. Mol. Biol.* **1994**, 243. 458-471

27 Laughlan, G., Murchie, A.I., Norman, D.G., Moore, M.H., Moody, P.C., Lilley, D.M. and Luisi, B. *Science* **1994**, 265, 520-524; Phillips, K., Dauter, Z., Murchie, A.I., Lilley, D.M. and Luisi, B. *J. Mol. Biol.* **1997**, 273, 171 –

182.

28 Wang, Y. and Patel, D.J. *Structure* **1994**, 2, 1141 – 1156.

29 Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. and Blackburn, E. H. *Cell* **1987**, 51, 899-908.

30 Williamson, J. R., Raghuraman, M. K. and Cech, T. R. *Cell* **1989**, 59, 871-880.

31 Sen, D. and Gilbert, W. *Nature* **1988,** 334, 364-366

32 Fang, G. and Cech, T. R. *Biochemistry*, **1993**, 32, 11646-11657.

33 Fang, G. and Cech, T. R. *Cell*, **1993**, 74, 875-885.

34 Giraldo, R. and Rhodes, D. *EMBO J.* **1994**, 13, 2411-2420.

35 Harrington, C., Lan, Y. and Akman, S. A. *J. Biol. Chem.* **1997**, 272, 24631-24636.

36 Lu, M., Guo, Q. and Kallenbach, N. R. *Biochemistry*, **1993**, 32, 598-601.

37 Wang, K. Y., Swaminathan, S. and Bolton, P. H. *Biochemistry*, **1994**, 33, 7517-7527.

38 Erlitzki, R. and Fry, M. *J. Biol. Chem.* **1997**, 272, 15881-15890.

39 Kool, E.T. *Accounts Chem. Res.* **1998**, 31, 502-510.

40 Kool, E.T. *Annu. Rev. Bioph. Biom.* **1996**, 25, 1-28.

41 Baran, N., Pucshansky, L., Marco, Y., Benjamin, S. and Manor, H. *Nucleic Acids Res.* **1997**, 25, 297-303.

42 Sun, H., Karow, J. K., Hickson, I. D. and Maizels, N. *J. Biol. Chem.* **1998**, 273, 27587-27592.

43 Arimondo, P. B., Riou, J-F., Mergny, J.-L., Tazi, J., Sun, J. S., Garestier, T. and Hélène, C. *Nucleic Acids Res.* **2000**, 28, 4832-4838.

44 Henderson E., Hardin, C. C., Walk, S. K., Tinoco, I. Jr. and Blackburn E. H.

Cell **1987**, 51, 899-908.

45 Veselkov, A. G., Malkov, V. A., Frank-Kamenetskll, M. D. and Dobrynin, V. N. *Nature* **1993**, 364, 496.

46 Sundquist, W. I. and Klug, A. *Nature* **1989**, 342, 825-829

47 Wellinger, R. J, Wolf, A. J. and Zakian, V. A. *Cell* **1993**, 72, 51-60.

48 Zahler, A. M, Williamson, J. R. and Cech, T. R., Prescott, D. M. *Nature* **1991**, 350, 718-720.

49 Mergny, J.L. and Helene, C. *Nat. Med.* **1998**, 4, 1366-1367.

50 Sharma, S., Raymond, E., Soda, H., Sun, D., Hilsenbeck, S. G., Sharma, A., Izbicka, E., Windle, B. and Von Hoff, D. D. *Ann. Oncol.* **1997**, 8, 1063-1074.

51 Perry, P. J. and Kelland, L.R. *Exp. Opin. Ther. Patents,* **1998**, 8, 1567-1586

52 Anantha, N. V., Azam, M. and Sheardy, R. D. *Biochemistry* **1998**, 37, 2709-2714.

53 Wheelhouse, R. T., Sun, D., Han, H., Han, F. X. and Hurley, L. H. *J. Am. Chem. Soc.* **1998**, 120, 3261-3262.

54 Han, F.X., Wheelhouse, R. T. and Hurley, L. H. *J. Am. Chem. Soc.* **1999**, 121, 3561-3570.

55 Haq I., Trent, J. O., Chowdhry, B. Z. and Jenkins, T. C. *J. Am. Chem. Soc.* **1999**, 121, 1768-1779.

56 Gehring, K. and Leroy, J.-L., GuJron, M. *Nature* **1993**, 363, 561 –564

57 Gueron, M.and Leroy, J.-L. *Curr. Opin. Struct. Biol.* **2000**, 10, 326 – 331

58 Kang, C., Berger, I., Lockshin, C., Ratliff, R. and Moyzis, R., Rich, A. *Proc. Natl. Acad. Sci. USA*, **1995**, 92, 3874 – 3878; Leroy, J.-L., Gehring, K., Kettani, A. and GuJron, M. *Biochemistry* **1993**, 32, 6019-6031; Voloshin, O. N., Veselkov, A. G., Belotserkovskii, B. P., Danilevskaya, O. N., Pavlova,

M. N., Dobrynin, V. N.and Frank-Kamenetskii, M. D. *J. Biomol. Struct. Dyn.* **1992**, 9, 643 –652;

59 Lacroix, L., Mergny, J.-L., Leroy, J.-L. and Hélène, C. *Biochemistry* **1996**, 35, 8715-8722; Mergny, J.-L., *Biochemistry* **1999**, 38, 1573-1581; Snoussi, K., Nonin-Lecomte, S. and Leroy, J.-L. *J. Mol. Biol.* **2001**, 309, 139-153; Gilbert, D. E. and Feigon, J., *Curr. Opin. Struct. Biol.* **1999**, 9, 305-314.

60 Fedoroff, O. Y., Rangan, A., Chemeris, V. V. and Hurley, L. H., *Biochemistry* **2000**, 39, 15083-15090; Mergny, J.-L., Lacroix, L., Han, X., Leroy, J.-L. and Hélène, C., *J. Am. Chem. Soc.* **1995**, 117, 8887-8898; Leroy, J. L., Gueron, M., Mergny, J. L. and Hélène, C., *Nucleic Acids Res.* **1994**, 22, 1600-1606; Gallego, J., Chou, S.-H. and Reid, B. R. *J. Mol. Biol.* **1997**, 273, 840-856.

61 Phan, A.T. and Leroy, J. - L. *J. Biomol. Struct. Dyn.* **2000**, 17, 377-383.

62 Patel, D. J., Bouaziz, S., Kettani, A. and Wang, Y. *Oxford Handbook of Nucleic Acid Structures.* Edited by Neidle S. New York, Oxford University Press, **1999**, 389-453.

63 Han, X., Leroy, J.-L.and Guéron, M. *J. Mol. Biol.* **1998**, 278, 949-965.

64 Huertas, D., Fanti, L., Pimpinelli, S., Marsellach, F. X., Pina, B., Azorin, F. *EMBO J* **1999**, 18, 3820-3833.

65 Lacroix, L., Lienard, H., Labourier, E., Djavaheri-Mergny, M., Lacoste, J., Leffers, H., Tazi, J., Helene, C. and Mergny, J. L. *Nucleic Acids Res.* **2000**, 28,1564-1575.

66 Mavroidis, C., Dubey, A. and Yarmush, M. L. *Annu. Rev. Biomed. Eng.* **2004**, 6, 10.1-10.33.

67 Ballardini, R., Balzani, V., Credl, A., Gandolfi, M. T. and Venturi, M. *Acc.*

Chem. Res **2001**, 34, 445-455.

68 P. D. Boyer, *Annu. Rev. Biochem.* **1997**, 66, 717-49.

69 Kojima, S. and Blair, D. F. *Int Rev Cytol.* **2004**, 233, 93-134.

70 Beissenhirtz, M. K. and Willner, I. *Org. Biomol. Chem.* **2006**, 4, 3392–3401.

71 Alberti, P., Bourdoncle, A., Sacca, B., Lacroix, L. and Mergny, J.-L. *Org. Biomol. Chem.* **2006**, 4, 3383–3391.

72 Simmel, F. C. and Dittmer, W. U. *Small* **2005**, 1, 284-299.

73 Condon. A. *Nature Reviews* **2006**, 7, 565.

74 Mao, C., Sun, W., Shen, Z. and Seeman, N. C. *Nature*, **1999**, 397, 144-146.

75 Yurke, B., Turberfield, A. J., Mills Jr, A. P., Simmel, F. C. and Neumann, J. L. *Nature*, 2000, 406, 605-609.

76 Simmel, F. C. and Yurke, B. *Phys. Rev.* **2001**, 63, 041913.

77 Mitchell, J. C. and Yurke, B. *DNA Scissors* Springer Berlin, Heidelberg, **2002,** Volume 2340, pp. 258-268

78 Shin, J. S., and Pierce, N. A. *J. Am.Chem. Soc.* **2004**, 126, 10834-10835; Yin, P., Yan, H., Daniell, X. G., Turberfield, A. J. and Reif, J. H. *Angew Chem Int Ed Engl.* **2004**, 43, 4906-11.

79 Phan, A.T. and Mergny, J. L. *Nucleic Acids Res.* **2002**, 30, 4618-4625.

80 Li, W., Miyoshi, D., Nakano, S. and Sugimoto, N. *Biochemistry.* **2003**, 42, 11736.

81 Risitano, A. and Fox, K. R. *Bioorg. Med. Chem. Lett.* **2005**, 15, 2047.

82 Alberti, P. and Mergny, J. L. *Proc. Natl. Acad. Sci. USA.* **2003**, 100, 1569-1573.

83 Dittmer, W. U., Reuter, A. and Simmel, F. C. *Angew. Chem Int. Ed*. **2004**, 43, 3550-3553; Dittmer, W. U. and Simmel, F. C. *Nano Letters* **2004**, 4, 689-691.

84 Bourdoncle, A., Estévez Torres, A., Gosse, C., Lacroix, L.,Vekhoff, P., Jullien, L. and Mergny J.-L. *J. Am. Soc.* **2006**, 128, 11094-11105.

85 Liedl, T. and Simmel, F. C. *Nano Letters* **2005**, 5, 1894-1898.

86 Liu, D. and Balasubramanian, S. *Angew. Chem. Intl. Ed.* **2003**, 42, 5734-5736.

87 Liu, D., Bruckbauer, A., Abell, C., Balasubramanian, S., Kang, D.-J., Klenerman, D. and Zhou, D. *J. Am. Chem. Soc*. **2006**, 128, 2067-2071.

88 Chen, Y., Wang, M. and Mao, C. *Angew. Chem Int. Ed*., **2004**, 43, 3554-3557; Chen, Y., Lee, S.H. and Mao, C. *Angew. Chem Int. Ed*., **2004**, 43, 53354-5338; Beyer, S., Dittmer, W. U. and Simmel, F. C. *J. Biomed. Nanotechnol*, **2005**, 1, 96-101; Simmel, F. C. and Yurke, B. *Appl. Phys. Lett.* **2002**, 80, 883-885.

89 Didenko, V. V., Minchew, C. L., Shuman, S. and Baskin, D. S. *Nano Letters* **2004**, 4, 2461-2466; Buranachai, C., McKinney, S. A. and Ha, T. *Nano Letters* **2006**, 6, 496-500; Viasnoff, V., Meller, A. and Isambert, H. *Nano Letters* **2006**, 6, 101-104; Li, J. J. and Tan, W. *Nano Letters* **2002**, 2, 315-318; Feng, L., Park, S. H., Reif, J. H. and Yan, H. *Angew. Chem. Int. Ed.* **2003,** 42**,** 4342-4346; Yan, H., Zhang, X., Shen, Z. and Seeman, N. C. *Nature* **2002**, 415, 62-65.

90 Betterton, M. D. and Julicher, F. *Phys. Rev.* **2005**, E 71, 011904.

91 Moore, K. J. M. and Lohman, T. M. *Biophys. J.* **1995**, 68, 180-185.

92 Gehring, K., Leroy, J. L.and Guéron, M. *Nature,* **1993**, 363, 561-565; Guéron, M. and Leroy, J. L. *Curr. Opin. Stru. Biol.* **2000**, 10**,** 326-331; Leroy, J. L. and Lacroix, L. *Nucleic Acids Res.* **1998**, 26, 4797-4803.

93 Karolin, J., Johansson, LB-Å., Strandberg, L. and Ny, T. *J. Am. Chem. Soc.*

1994, 116, 7801-7806.

94 Haugland, R.P., *Optical Microscopy for Biology*, Herman. B., Jacobson, K., Eds. **1990**, pp. 143-157.

95 Vos de Wal, E., Pardoen, J.,van Koeveringe, J. A. and Lugtenburg, *J. Recl. Trav. Chim. Pays-Bas* **1977**, 96, 306-309.

96 Okamura, Y., Kondo, S., Sase, I., Suga, T., Mise, K., Furusawa, I., Kawakami, S. and Watanabe, Y. *Nucleic Acids Res.* **2000**, 28, e107; Sei-Iida, Y., Koshimoto, H., Kondo, S. and Tsuji, A. *Nucleic Acids Res*. **2000**, 28, e59.

97 Tyagi, S. and Kramer, F. R. *Nat. Biotechnol*. **1996**, 14, 303-308; Niemeyer, C. M and Adler, M. *Angew. Chem Int. Ed*. **2002**, 41, 3779-3783.

98 Staschke, K. A., Richardson, K. K., Mabry, T. E., Baxter, A. J., Scheuring, J. C., Huffman, D. M., Smith, W. C., Richardson, F. C. and Colacino, J. M. *Nucleic Acids Res*. **1996,** 24, 4111-4116.

99 Shu, W., Liu, D., Watari, M., Strunz, T., Riener, C., Welland, M. E., Balasubramanian, S. and McKendry, R. A. *J. Am. Chem. Soc.* **2005**, 127, 17054-17060.

100 Marky, L. A. and Breslauer, K. J. *Biopolymers* **1987**, 26, 1601-1620

101 Breaker, R. R. and Joyce, G. F. *Chem Biol.* **1994**, 1, 223-229.

102 Silverman, S. K. *Org. Biomol. Chem.* **2004**, 2, 2701-06.

103 Liu, J. and Lu, Y. *Chem. Mater.* **2004**, 16, 3231-38.

104 Li, Y., Sen, D. *Nat. Struct. Biol.* **1996**, 3, 743; Roth, A., Breaker, R. R. *Proc. Natl. Acad. Sci. U. S.* **1998**, 95, 6027.

105 Breaker, R. R. and Joyce, G. F. *Chem. & Biol*. **1994**, 1, 223-229; Cuenoud, B. and Szostak, W. J. *Nature* **1995**, 375, 611-614; Li, Y. and Sen, D. *Nat. Struct. Biol.* **1996**, 3, 743-747.

106 Faulhammer, D. and Famulok, M. *Angew. Chem., Int. Ed. Engl.* **1996**, 35, 2837-2841; Li, Y. and Breaker, R.R. *Proc. Natl. Acad. Sci. U. S.* **1999**, 96, 2746-2751; Li, Y., Liu, Y. and Breaker, R.R. *Biochem.* **2000**, 39, 3106-3114.

107 Carmi, N. and Breaker, R.R. *Bioorg. Med. Chem.* **2001**, 9, 2589-2600; Sreedhara, A., Li, Y. and Breaker, R.R. *J. Am. Chem. Soc.* **2004**, 126, 3454-3460; Chinnapen, D. J. F. and Sen, D. *Proc. Natl. Acad. Sci.*, **2004**, 101, 65-69.

108 Liu, X., Li, X., Zhou, T., Wang, Y., Ng, M. T. T., Xu, W. and Li, T. *Chem.Commun.* **2008**, 380–382

109 Simonsson, T. *Biol. Chem.* **2001**, 382, 621-628.

110 Miyoshia, D., Nakaoa, A., Todaa, T. and Sugimotoa, N. *FEBS Letters* **2001**, 496, 128-133; Miyoshi, D., Nakao, A. and Sugimoto, N. *Nucleic Acids Research* **2003**, 31, 1156-1163; Chen, F.-M. *Biochemistry* **1992**, 31, 3769-3776; Smirnov, I. and Shafer, R. H. *J. Mol. Biol.* **2000**, 296, 1-5.

111 Browning, D. D., Mc, S.M., Marty, C. and Ye, R. D. *J. biol. Chem.* **2001**, 276, 13039-13048.

112 Adkins, T. S., Perrotta, A.T., Ferré-D'Amaré, A. R., Doudna, J. A. and Been, M.D. *RNA* **1999**, 5, 720- 726.

113 Iter, N. G. and Burker, J. M. *Current Opinion in Chemical Biology* **1998**, 2, 24-30.

114 Carmi, N., Shultz, L.A. and Breaker, R. R. *Chem Biol.* **1996**, 3, 1039-1046.

115 Breaker, R. R. and Joyce, G. F. *Chemistry & Biology*, **1995**, 2, 655-660.

116 Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. and Neidle, S. *Nucleic Acids Res.* **2006**, 34, 5402-5415; Simonsson, T. *Biol. Chem.* **2001**, 382, 621.

117 Belmont, P., Constant, J. F. and Demeunynck, M. *Chem. Soc. Rev.* **2001**, 30,

70-81; Lane, A.N. and Jenkins, T. C. *Curr. Org. Chem.* **2001**, 5, 845-869; Schultze, P., Hud, N. V., Smith, F. W. and Feigon, J. *Nucleic Acids Res.* **1999**, 27, 3018; Pilch, D. S., Plum, G. E. and Breslauer, K.J. *Curr. Opin. Struct. Biol.* **1995**, 5, 334-342.

118 Phan, A. T., Kuryavyi, V. and Patel, D. J. *Curr. Opin. Struct. Biol.* **2006**, 16, 288-298; Krishnan-Ghosh, Y., Liu, D. andBalasubramanian, S. *J. Am. Chem. Soc*. **2004**, 126, 11009-11016; Parkinson, G. N., Lee, M. P. H. and Neidle, S. *Nature* **2002**, 417, 876-880.

119 Dolinnaya, N. G., Blumenfeld, M., Merenkova, I. N., Oretskaya, T. S., Krynetskaya, N. F., Ivanovskaya, M. G., Vasseur, M. and Shabarova, Z. A. *Nucleic Acids Res.* **1993**, 21, 5403-5407.

120 Wang, S. and Kool, E. T. *Nucleic Acids Res.* **1994**, 22, 2326-2333.

121 Rujan, N. I., Meleney, J. C. and Bolton, P. H. *Nucleic Acids Res.* **2005**, 33, 2022-2031; Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. and Neidle, S. *Nucleic Acids Res.* **2006**, 34, 5402-5415.

122 Williamson, J. R. *Annu. Rev. Bioph. Biom.*, **1994**, 23, 703.

123 Simonsson, T. *Biol. Chem*. **2001,** 382, 621.

124 Smith, F. W. and Feigon, J. *Nature* **1992**, 356, 164-168.

125 Jin, R., Breslauer, K. J., Jones, R. A. and Gaffney, B. L. *Science* **1990**, 250, 543-546.

126 Jin R., Gaffney, B. L., Wang, C., Jones, R. A. and Breslauer K. J. *Proc. Natl. Acad. Sci. USA* **1992**, 89, 8832-8836.

127 Marathias, V. M. and Bolton, P. H. *Nucleic Acids Res.* **2000**, 28, 1969-1977.

128 Scaria, P. C., Shire, S. J. and Shafer, R. H. *Proc. Natl. Acad. Sci. USA* **1992**, 89, 10336-10340.

129 Blume, S. W., Guarcello, V., Zacharias, W. and Miller, D. M. *Nucleic Acids Res.* **1997**, 25, 617-625.

130 Murchie, A. I. and Lilley, D. M. *Nucleic Acids Res.* **1992**, 20, 49-53.

131 Tuntiwechapikul, W., Lee, J. T. and Salazar, M. *J. Am. Chem. Soc.* **2001**, 123, 5606-5607.

132 Wang, Y. and Patel, D. J. *Structure* **1993**, 1, 263-282.

133 Lu, M., Guo, Q. and Kallenbach, N. R. *Biochemistry* **1993**, 32, 598-601.

134 H. Deng, and Braunlin, W. H. *J. Mol. Biol.* **1996**, 255, 476-483.

135 Kankia, B. I. and Marky, L. A. *J. Am. Chem. Soc.* **2001**, 123, 10799-10804.

136 Han, H., Langley, D. R., Rangan, A. and Hurley, L. H. *J. Am. Chem. Soc.* **2001**, 123, 8902-8913.

137 Smith, S. S., Baker, D. J. and Jardines, L. A. *Biochem. Biophys. Res. Commun.* **1989**, 160, 1397-1402.

138 Wang, K. Y., Swaminathan, S. and Bolton, P. H. Biochemistry **1994**, 33, 7517-7527.

139 *Schultze, P., Macaya, R. F. and Feigon, J.* J. Mol. Biol. ***1994***, *235, 1532-1547.*

140 Padmanabhan, K., Padmanabhan, K. P., Ferrara, J. D., Sadler, J. E. and Tulinsky, A. *J. Biol. Chem.* **1993**, 268, 17651-17654.

141 Kang, C., Zhang, X., Ratliff, R., Moyzis, R. and Rich, A. *Nature* **1992**, 356, 126-131.

142 Schultze, P., Smith, F. W. and Feigon, J. *Structure* **1994**, 2, 221-233.

143 Laughlan, G., Murchie, A. I. H., Norman, D. G., Moore, M. H., Moody, P. C. E.; Lilley, D. M. and Luisi, B. *Science* **1994**, 265, 520-524.

144 Aboul-ela, F., Murchie, A. I. and Lilley, D. M. *Nature* **1992**, 360, 280-282.

145 Chen, J., Liu, D., Lee, A. H. F., Qi, J., Chan, A. S. C.and Li, T. *Chem. Comm.* **2002**, 22, 2686-2687.

146 Jing, N., Marchand, C. Liu, J., Mitra, R., Hogan, M. E. and Pommier, Y. *J. Biol. Chem.* **2000,** 275, 21460-21467.

147 Macaya, R. F., Schultze, P., Smith, F. W., Roe, J. A. and Feigon, J. *Proc. Natl. Acad. Sci. USA* **1993**, 90, 3745-3749.

148 Erlitzki, R. and Fry, M. *Biol. Chem.* **1997**, 272, 15881-15890.

149 Williamson, J. R., Raghuraman, M. K. and Cech T. R. *Cell* **1989**, 59, 871-880.

150 Simonsson, T., Pecinka, P. and Kubista, M. *Nucleic Acids Res.* **1998**, 26, 1167-1172.

151 Hammond-Kosack, M. C., Kilpatrick, M. W. and Docherty, K. *J. Mol. Endocrinol*. **1992**, 9, 221-225.

152 Kool, E.T. *Accounts Chem. Res*. **1998**, 31, 502-510.

153 Kool, Eric T. *Annu. Rev. Bioph. Biom*. **1996**, 25, 1-28.

154 Fang, G. and Cech, T. R. *Biochemistry***1993**, 32, 11646-11657.

155 Fang, G. and Cech, T. R. *Cell* **1993**, 74, 875-885.

156 Giraldo, R. and Rhodes, D. *EMBO J.* **1994**, 13, 2411-2420.

157 Harrington, C., Lan, Y. and Akman, S. A. *J. Biol. Chem.* **1997**, 272, 24631-24636.

158 Baran, N., Pucshansky, L., Marco, Y., Benjamin, S. and Manor, H. *Nucleic Acids Res.* **1997**, 25, 297-303

159 Sun, H., Karow, J. K., Hickson, I. D. and Maizels, N. *J. Biol. Chem.* **1998**, 273, 27587-27592.

160 Fry, M. and Loeb, L. A. *J. Biol. Chem.* **1999**, 274, 12797-12802.

161 Browning, D. D., Mc, S.M., Marty, C. and Ye, R. D. *J. biol. Chem.* 2001, 276, 13039-13048.

162 Wehrl, W., Niederweis, M. and Schumann, W. *J. Bacteriol.* **2000**, 182, 3870-3873.

163 Patel, D. J., Bouaziz, S., Kettani, A. and Wang, Y. In Neidle, S. (Ed.), *Oxford Handb. Nucleic Acid Struct.*, Oxford University Press, Oxford, UK, **1999**, pp. 389-453.

164 Wang, S. and Kool, E. T. *Nucleic Acids Res.* **1994**, 22, 2326-2333.

165 Clark, G. R., Pytel, P. D., Squire, C. J. and Neidle, S. *J. Am. Chem. Soc.* **2003**, 125, 4066-4067.

166 Haider, S. M., Parkinson, G. N. and Neidle, S. *J. Mol. Biol.* **2003**, 326(1), 117-125.

167 Kang, Sang-G. and Henderson, E. *Mol. Cells* **2002**, 14, 404-410.

168 Lyonnais, S., Gorelick, R. J, Mergny , Jean-L., Le, C.E. and Mirambeau, G. *Nucleic Acids Res.* **2003**, 31, 5754-63.

169 Hurley, L. H. *Biochem. Soc. T.* **2001**, 29, 692-696.

170 Wen, Jin-D. and Gray, D. M. *Biochemistry* **2002**, 41, 11438-11448.

171 Ying, L., Green, J. J., Li, H., Klenerman, D. and Balasubramanian, S. *Proc. Natl Acad. Sci. USA* **2003**, 100, 14629-14634.

172 Kim, M-Y., Gleason-Guzman, M., Izbicka, E., Nishioka, D. and Hurley, L. H. *Cancer Res.* **2003**, 63, 3247-3256.

173 Siddiqui-Jain, A., Grand, C. L., Bearss, D. J. and Hurley, L. H. *Proc. Natl Acad. Sci. USA* **2002**, 99, 11593-11598.

174 Dolinnaya, N. G., Blumenfeld, M., Merenkova, I. N., Oretskaya, T. S., Krynetskaya, N. F., Ivanovskaya, M. G., Vasseur, M. and Shabarova, Z. A.

Nucleic Acids Res. **1993**, 21, 5403-5407.

175 Simonsson, T. *Biol. Chem.* **2001**, 382, 621-628.

176 Williamson, J. R. *Annu. Rev. Bioph. Biom*. **1994**, 23, 703-730

177 Balagurumoorthy, P., Brahmachari, S. K., Mohanty, D., Bansal, M. and Sasisekharan, V. *Nucleic Acids Res.* **1992,** 20, 4061-4067.

178 Guo, Q., Lu, M. and Kallenbach, N. R. *Biochemistry* **1993**, 32, 3596-3603.

179 Smirnov, I. and Shafer, R. H. *Biochemistry* **2000**, 39, 1462-1468.

180 Kool, E. T. *Acc. Chem. Res.* **1998**, 31, 502.

181 Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **1996**, 25, 1

182 Wang, S. and Kool, E. T. *Nucleic Acids Res.* **1994**, 22, 2326.

183 Clark, G. R., Pytel, P. D., Squire, C. J. and Neidle, S. *J. Am. Chem. Soc.* **2003**, 125, 4066.

184 Mergny, J. L. and Maurizot, J. C. *J. Am. Chem. Soc.* **2003**, 125, 4066.

185 Leppard, J. B. and Champoux, J. J. *Chromosoma* **2005**, 114, 75-85.

186 Champoux, J. J. *Annu. Rev. Biochem.* **2001**, 70, 369-413.

187 Wang, J. C. *Nat. Rev. Mol. Cell Biol.* **2002**, 3, 430; Corbett K. D. and Berger, J. M. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, 33, 95.

188 Giovanella, B. C., Stehlin, J. S., Wall, M. E., Wani, M. C., Nicholas, A. W., Liu, L. F., Silber, R.and Potmesil, M. *Science* **1989**, 246, 1046; Husain, I., Mohler, J. L., Seigler, H. F. and Besterman, J. M. *Cancer Res.* **1994**, 54, 539.

189 Van der Zee, A. G. J., Hollema, H., De Jong, S., Boonstra, H., Gouw, A., Willemse, P. H. B., Zijlstra, J. G. and De Vries, E. G. E. *Cancer Res.* **1991**, 51, 5915; Hirabayashi, N., Kim, R. and Nishiyama, M. *Proc. Am. Assoc. Cancer Res.* **1992**, 33, 436.

190 Pommier, Y. *Nat. Rev. Cancer* **2006**, 6, 789; Topcu, Z. J. *Clin. Pharm. Ther.*

2001, 26, 405; Rothenberg, M. L. *Ann. Oncol.* **1997**, 8, 837; Li T. K. and Liu, L. F. *Annu. Rev. Pharm. Toxicol.* **2001**, 41, 53; Hertzberg, R. .P., Caranfa, M. J. and Hecht, S. M. *Biochemistry*, **1989**, 28, 4629.

191 Mortensen, U. H., Stevnsner, T., Krogh, S., Olesen, K., Westergaard, O. and Bonven, B. J. *Nucleic Acids Res.*, **1990**, 18, 1983; Li, C. J., Averboukh, L. and Pardee, A.B. *J. Biol. Chem.*,**1993**, 268, 22463.

192 Edwards, K. A., Halligan, B. D., Davis, J. L., Nivera, N. L. and Liu, L. F. *Nuclei Acids Res*. **1982**, 10, 2565-2576

193 Been, M. D., Burgress, R. R. and Champoux, J. J. *Nuclei Acids Res*. **1984**, 12, 3097- 3114

194 Bonven, B. J., Gocke, E. and Westergaard, O. *Cell* **1985**, 41, 541-551

195 Christiansen, K., Bonven, B. J. and Westergaard, O. *J. Mol. Biol*. **1987**, 193, 517-525

196 Ness, P. J., Koller, T. and Thoma, F. *J. Mol. Biol*. **1988**, 200, 127-139

197 Svejstrup, J. Q., Christiansen, K., Andersen, A. H., Lund, K. and Westergaard, O. *J. Biol. Chem*. **1990**, 265, 12529- 12535

198 Li, X., Ng, M. T. T., Wang, Y., Liu X. and Li, T. *Bioorg. Med. Chem. Lett.*, **2007**, 17, 4967.

199 Moody, E. M. and Bevilacqua, P. C. *J. Am. Chem. Soc.* **2003**, 125, 16285; Moody, E. M. and Bevilacqua, P. C. *J. Am. Chem. Soc.*, **2003**, 125, 2033; Kuch, D., Schermelleh, L., Manetto, S., Leonhardt, H. and Carell, T. *Angew.Chem.Int.Ed.* **2008**, 47, 1515.

200 Yamakawa, H., Abe, T., Saito, T., Takai, K., Yamamoto, N. and Takaku, H. *Bioorg. Med. Chem.*, **1998**, 6, 1025; Owczarzy, R., Vallone, P. M., Goldstein, R. F. and Benight, A. S. *Biopolymers*, **1999**, 52, 29; Clusel, C.,

Ugarte, E., Enjolras, N., Vasseur, M. and Blumenfeld, M. *Nucleic Acids Res.* **1993**, 21, 3405.

201 Yoshizawa, S., Ueda, T., Ishido, Y., Miura, K., Watanabe, K. and Hirao, I. *Nucleic Acids Res*. **1994**, 22**,** 2217.

202 Leppard, J. B. and Champoux, J. J. *Chromosoma* **2005**, 114, 75; Roca, J. *Trends Biochem. Sci.,* **1995**, 20**,** 156.

203 Christiansen, K., Svejstrup, A. B., Andersen, A. H.and Westergaard, O. J. *Biol. Chem.,* **1993**, 268, 9690; Sikder, D. and Nagaraja, V. *J. Mol. Biol.*, **2001**, 312, 347; Svejstrup, J. Q., Christiansen, K., Andersen, A. H., Lund, K. and Westergaard, O. *J. Biol. Chem.,* **1990**, 265, 12529.

204 Pourquier, P., Ueng, L. M., Kohlhagen, G., Mazumder, A., Gupta, M., Kohn, K. W. and Pommier, Y. *J. Biol. Chem.*, **1997**, 272, 7792; Henningfeld, K. A. and Hecht, S. M. *Biochemistry,* **1995**, 34, 6120; Yeh, Y. C., Liu, H. F., Ellis, C. A. and Lu, A. L. *J. Biol. Chem.*, **1994**, 269, 15498.

205 Antony, S., Arimondo, P. B., Sun, J. S. and Pommer, Y. *Nucleic Acids Res*., **2004**, 32, 5163; Wang, X., Henningfeld, K. A. and Hecht, S. M. *Biochemistry* **1998**, 37, 2691; Straub, T., Knudsen, B. R. and Boege, F. *Biochemistry* **2000**, 39, 7552; Svejstrup, J. Q., Andersen, A. H. and Westergaard, O. *J. Mol. Biol.*, **1991**, 222, 669.

206 Laine, J. P., Opresko, P. L., Indig, F. E. , Harrigan, J. A., Kobbe, C. V. and Bohr, V. A. *Cancer Res.,* **2003**, 63, 7136.

207 Modrich, P. *J. Biol. Chem.*, **2006**, 281, 30305; Bohr, V.A., Wassermann, K. and Kraemer, K. H. *DNA Repair Mechanisms,* Alfred Benzon Symposium, Copenhagen, Munksgaard, **1993**, 35**,** pp.1-428.

208 Xu, M.-Q. and Evans, T. C. *Methods* **2001,** 24, 257; Parkinson, M J.

and Lilley, D. M. J. *J. Mol. Biol.*, **1997**, 270, 169.

209 http://www.neb.com/nebecomm/products/productM0201.asp

210 http://www.neb.com/nebecomm/products/productM0202.asp

211 http://www.neb.com/nebecomm/products/productM0293.asp

212 http://www1.gelifesciences.com/APTRIX/upp01077.nsf/Content/Products?OpenDocument&parentid=41359&moduleid=41368#content

213 http://www.neb.com/nebecomm/products/productM0303.asp

214 http://www.neb.com/nebecomm/products/productM0302.asp

215 http://www.neb.com/nebecomm/products/productM0263.asp

216 Kerr, C. and Sadowski, P. D. *J. Biol. Chem.* **1972**, 247, 305-318.

217 Shinozaki, K. and Okazaki, T. *Nucl. Acids Res.* **1978**, 5, 4245-4261.

218 http://www.neb.com/nebecomm/products/productR0525.asp

219 Morgan, R. D. *Nucleic Acids Res.* **1988**, 16, 3104.

220 http://www.topogen.com/html/recombinant_human_topo_i.html

221 http://www.neb.com/nebecomm/products/productN3033.asp

222 Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M.C., Heynecker, H.L. and Boyer, H.W. *Gene* **1977**, 2, 95-113.

223 Watson, N. *Gene* **1988**, 70, 399-403.

224 *Molecular cloning: A laboratory manual. (Second Edition)* Edt. Maniatis, T.; Fritsch, E.F.; Sambrook, J. Cold Spring Harbor Laboratory, Cold Spring Harbor, **1987,** N. Y.

225 Luebke, K. J. and Dervan, P.B. *Nucl. Acids Res.* **1992**, 20, 3005-3009.

List of Publications

1 Yifan Wang, Magdeline Tao Tao Ng, Tianyan Zhou, Choon Hong Tan and Tianhu Li. C3-spacer-containing circular oligonucleotides as inhibitors of human topoisomerase I. *Bioorganic & Medicinal Chemistry Letters*, **2008**, 18, 3597-3602.

2 Yifan Wang, Xinming Li, Xiaoqian Liu and Tianhu Li. An i-motif-containing DNA Device that Breaks Certain Forms of Watson-Crick Interactions. *Chemical Communications*, **2007**, 42, 4369 – 4371

3 Yifan Wang, Wei Xu, Tianyan Zhou, Xinming Li, Magdeline Tao Tao Ng, Xiaoqian Liu, Sock Teng Chua and Tianhu Li. Studies of factors that affect self-cleaving reactions of G-quadruplexes. (*submitted)*

4 Tianyan Zhou, Yifan Wang, Xinming Li, Qiang Zhang and Tianhu Li. Synthesis and Characterization of a Fluorescein-Labeled Circular G-Quadruplex. *Bulletin of the Chemical Society of Japan,* **2006**, Vol. 79, No. 8, pp.1300-1302

5 Xinming Li, Magdeline Tao Tao Ng, Yifan Wang, Tianyan Zhou, Sock Teng Chua, Weixing Yuan and Tianhu Li. Site-specific self-cleavage of G-quadruplexes formed by human telomeric repeats. *Bioorganic & Medicinal Chemistry Letters*, **2008**, 18, 5576-5580

6 Xiaoqian Liu, Xinming Li, Tianyan Zhou, Yifan Wang, Magdeline Tao Tao Ng , Wei Xu and Tianhu Li. Site specific self-cleavage of certain assemblies of G-quadruplex. *Chemical Communications,* **2008**, 3, 380-382

7 Xinming Li, Magdeline Tao Tao Ng, Yifan Wang, Xiaoqian Liu and Tianhu Li. Dumbbell-shaped circular oligonucleotides as human topoisomerase I inhibitors. *Bioorganic & Medicinal Chemistry Letters*, **2007**, 17, 4967–4971

8 Tianyan Zhou, Guoshu Chen, Yifan Wang, Qiang Zhang, Ming Yang and Tianhu Li. Synthesis of unimolecularly circular G-quadruplexes as prospective molecular probes. *Nucleic Acids Research*, **2004**, 32(21), e173/1-e173/9.

CONFERENCE PAPERS/POSTER PRESENTATION:

1 Yifan Wang and Tianhu Li, Synthesis of Dumbbell-Shaped Circular Oligonuleotides Containing Internal C3-Spacers as Human Topoisomerase I Inhibitors. Singapore International Chemistry Conference 5 (*SICC-5*) & 7th Asia-Pacific International Symposium on Microscale Separation and Analysis (*APCE 2007*), December 16-19, **2007**, Suntec Convention & Exhibition Centre, Singapore

2 Yifan Wang and Tianhu Li, A new energy-converting DNA machine driven by the formation and dissociation of i-motif. First International Meeting on Quadruplex DNA, April 21-24, **2007**, The Brown Hotel, Louisville, KY, USA

3 Yifan Wang and Tianhu Li, Synthesis of Fluorescein-Tagged Circular i-Motif as Prospective Molecular Probes. 9th international symposium for Chinese Organic Chemists (ISCOC-9), December 17-21 **2006**, Grand Copthorne Waterfront Hotel, Singapore

4 Yifan Wang and Tianhu Li, Synthesis of Circular G-quadruplex Tagged with Fluorescein. First Singapore Mini-Symposium on Medicinal Chemistry "Advances in Synthesis and Screening", July 6, **2005**, National University of Singapore, Singapore